Understanding the genetic and physiology controls of 'crumbly' fruit in red raspberry (*Rubus idaeus*).

Luca Maria Scolari

A thesis submitted for the degree of

Doctor of Philosophy

School of Engineering and Physical Sciences

December 2020

Abstract

'Crumbly' fruit is a generic term to indicate a phenomenon affecting raspberry (*Rubus idaeus*) and causing the formation of misshapen fruits characterised, in general, by very large drupelets but drastically reduced in number. The imperfect adhesion between these abnormal drupelets causes the fruit to crumble when is picked up and hence the name 'crumbly' fruit.

Symptoms vary in degree, with very severe forms, genetically determined and here defined as Crumbly Fruit Condition (**CFC**), to more variable and environmental related phenotypes here defined as Malformed Fruit Disorder (**MFD**). 'Crumbly' fruit can cause high yield losses and in particular **MFD**, studied in this thesis, due to its unpredictable expression across different seasons is becoming a serious threat for the raspberry industry.

Early studies stated that ''crumbly' fruit is an indication of a partial failure in one or more physiological processes concerned with fruit development'. The aim of this project was to study, in red raspberry, the physiology and the molecular processes behind fruit development to help control or eradicate the phenomenon.

'Crumbly' fruit phenotypic data, over many seasons, from a population of Latham x Glen Moy were re-analysed on a new Genotype by Sequencing (GbS) high density linkage map. The analysis identified a new 'crumbly' quantitative trait locus (QTL) on linkage group 3 (LG3) and confirmed the two previously identified QTL on LG1 and LG3. From the same population, transcriptomic analysis, via RNA microarray experiments identified genes differently expressed, some of which mapped inside the 'crumbly' QTLs. The study of the function of these differently expressed genes showed how impairments in processes related to pollen formation, pollen tube elongation and its interaction with the ovule might be responsible for the occurrence of this phenomenon.

Target phytohormones analyses on artificially induced 'crumbly' fruits at two different stages (i.e. green and red berry) and in two different parts (i.e. drupelet and receptacle) showed statistically significant differences in abscisic acid (ABA) and in gibberellin A1 (GA_1) levels, respectively in receptacle at green stage and in drupelet at red stage. These findings indicated the important role played by ABA and GA_1 in 'crumbly' fruit and in particular of ABA, as confirmed by the significant difference in the expression levels of numerous genes abscisic acid related (e.g. activated in response to ABA or related to its metabolism, biosynthesis and signalling pathway) found in the transcriptomic analysis.

I dedicated this work to perseverance, my loyal friend along these three years.

I do so with an aphorism in my native dialect (i.e. Neapolitan). It is an allegory to being patient and calmly waiting until the outcome of your actions is gotten.

"Dicette 'o pappice vicino 'a noce, damme 'o tiempo ca te spertose"

Acknowledgements

Firstly, I acknowledge my two sponsors, the Agricultural and Horticultural Development Board (AHDB) and the Mylnefield Trust. Without their financial support nothing about this work could have been never made.

I thank and acknowledge all my supervisors: Prof. Derek Stewart, Dr. Julie Graham, Dr. Rob Hancock, Dr. Craig Simpson, Dr. Alison Dolan and Nikki Jennings. The passion, the enthusiasm, the professionalism are distinguishing features to all of them and to me was very inspiring. It was with their support and guidance that I was on top of everything during these three years and a half of my PhD. I want to thank and acknowledge too, for the guidance, my two sponsor's supervisors Dr. Rachel McGauley and Dr. Louise Sutherland and so do I for my two liaison supervisors Dr. Sue Jones and Dr Alexandre Foito.

I want also to thank and acknowledge the help and the support I received from all the soft fruit group and from the staff in the LC-MS facility at The James Hutton Institute. So, my sincerest thanks to: Kay Smith, Linzi Jorgensen, Avril Britten, Amanda Rosa Moreno Mellado, Dr. Dorota Jarret, Dr. Dominique Williams, John Fuller, Sara Robertson, Dr. William Allwood and Dr. Sabine Freitag.

A special thanks to all the field and glasshouse stuff, to the postgraduate office of both Heriot-Watt university and The James Hutton Institute and in general to all the people that directly or indirectly contributed to the accomplishment of this work.

Last but not least, I want to acknowledge my family and in particular my parents Franco and Agata, my siblings and my wee niece Rebecca. They kept me feeling at home despite the distance; their moral support was priceless.

Research Thesis Submission

Name*:*	Luca Maria Scolari		
School:	Engineering and Physical Sciences		
Version: *(i.e. First, Resubmission, Final)*	Final	Degree Sought:	Doctor of Philosophy

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1. The thesis embodies the results of my own work and has been composed by myself
2. Where appropriate, I have made acknowledgement of the work of others
3. The thesis is the correct version for submission and is the same version as any electronic versions submitted*.
4. My thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require
5. I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.
6. I confirm that the thesis has been verified against plagiarism via an approved plagiarism detection application e.g. Turnitin.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate*:*	Luca Maria Scolari	Date:	03/12/2020

Submission

Submitted By *(name in capitals):*	LUCA MARIA SCOLARI
Signature of Individual Submitting:	Luca Maria Scolari
Date Submitted:	03/12/2020

For Completion in the Student Service Centre (SSC)

Limited Access	Requested	Yes		No		Approved	Yes		No
1.1 E-thesis Submitted (mandatory for final theses)									
Received in the SSC by *(name in capitals):*				Date:					

Tables of contents

List of Tables

Tables in Appendix

List of figures

Figure in appendix

Abbreviations

µg/µL	microgram/microliter
ABA	abscisic acid
AUX	auxin
CFC	Crumbly Fruit Condition
CK	cytokinin
cM	centimorgan
cv	cultivar
DNA	deoxyribonucleic acid
DNAse	deoxyribonuclease
FC	floricane
GA	Gibberellic acid
GA_1	Gibberellic acid 1
GA_4	Gibberellic acid 4
GbS	Genotype by Sequencing
GWAS	Genome Wide Association Study
ha	hectar
HPLC	High Performance Liquid Chromatography
IAA	Indole-3-acetic acid
LG	Linkage group
LOD	Logarithm of the odds
LoD	Limit of Detection
LoQ	Limit of Quantification
LT	low temperature
MFD	Malformed Fruit Disorder
MRM	Multiple Reaction Monitoring
m/z	mass/charge ratio
ng	nanogram
PCR	Polymerase Chain Reaction
PHCS	Plant Health Certification Scheme
PF	primocane

ppm	part per million
QqQ-LC-MS/MS	Triple Quadrupole Liquid Chromatography Mass Spectrometer
QTL	Quantitative Trait Loci
RBDV	Raspberry Bush Dwarf Virus
RLMV	Raspberry Leaf Mottle Virus
RNA	ribonucleic acid
RNAse	ribonuclease
RpLV	Raspberry Latent Virus
SA	salicylic acid
SD	short day
SNP	Single Nucleotide Polymorphism
SPE	Solid Phase Extraction
SSR	Simple Sequence Repeat

Chapter 1: General introduction

1.1 The red raspberry

Demand for raspberries both fresh and processed is at an all-time high with production rapidly increasing year on year (Strik et al., 2007), mainly fuelled by their perceived health benefits coupled with their pleasant taste. The fact that *Rubus* berries have some of the highest levels of antioxidants and phytonutrients of any fruit crop, due primarily to their intense concentration of anthocyanins and phenolic compounds has led to a number of investigations on antioxidant levels of raspberries (Anttonen and Karjalainen, 2005, Beekwilder et al., 2005, Dossett et al., 2008, Moore et al., 2008, Moyer et al., 2002). These have shown that the unique phytochemical profile of raspberry lends itself to a dietary route to health benefits, e.g. good gut health, type II diabetes, neuroprotections etc (Hancock et al., 2018, Kshatriya et al., 2019, Mukherjee and Ahmad, 2018, Garcia-Mazcorro et al., 2018).

Most raspberry production (e.g. 54% of global production updated 2017) is concentrated in northern and central European countries although there is an increasing interest in growing cane fruits in southern Europe (e.g. in Greece, Italy, Portugal and Spain). In many production areas, the fruit is grown primarily for the fresh market or high value juice products but in central Europe (e.g. Poland, Hungary and Serbia), a high proportion of the crop is destined for processing. North America (i.e. Canada, Mexico and USA) accounts for 29% of global production (https://www.tridge.com). Major regions of production in North America include the Pacific North West, California, Texas and Arkansas, as well as regions in New York, Michigan, Pennsylvania and Ohio which account for 13% world production. Chile, Argentina and Guatemala also have extensive production. Russia produces 18% of global production. An estimation of the worldwide production areas (93,229 ha) and annual production (612,570 tons) for the year 2014 was published by the Food and Agriculture Organization Corporate Statistical (FAOSTAT) Database (Linck and Reineke, 2019). In 2017 this rose to 795,250 tons (https://www.tridge.com). Although Europe, Russia and North America are still the main areas of raspberry production, in Asia and particularly in China, in the last few years a sharp increase in the area dedicated to this crop has been registered; other emerging countries in the production of red raspberry are in Oceania, New Zealand and temperate regions of Australia while in Africa the main players in raspberry industry are: Algeria, Kenya, Morocco and South Africa (Graham and Brennan, 2018b).

In theory, under ideal condition, a raspberry plantation can continue to fruit for more than 15 years but in practise the lifespan is drastically reduced due, mainly, to pathogen and pest threats (Jennings, 1988). Therefore, the selection of varieties bearing traits that give resistance to pathogens and pests, together with agroecological practices represent an important strategy to prolong the duration of plantations (Graham and Brennan, 2018).

Raspberries belong to the genus *Rubus* in the *Rosaceae* family. This comprises a high number of species, 250-700 depending on classification (Thompson, 1997). The genomic number is 7 and ploidy levels range from diploid to dodecaploid (Jennings, 1988, Meng and Finn, 2002). Members can be difficult to classify into distinct species due to hybridization and apomixes (Robertson, 1974), but molecular studies of phylogeny (Alice and Campbell, 1999, Sochor et al., 2015, Wang et al., 2016) are assisting the development of a robust phylogeny for *Rubus*. Species are distributed widely across Asia, Europe, North and South America, with the centre of diversity now considered to be in China. For a brief history of the crop from ancient times see Kellogg et al. (2011).

Details of the growth cycle have been described in detail by Jennings (1988). In red raspberry, according to plant life cycle, two different groups of cultivars are distinguished. The biennial fruiting varieties (2-year life cycle) present canes concluding all their life phases (i.e. vegetative growth, flower formation, fruiting as well as bud dormancy induction and its breaking) in two years; those biennial fruiting raspberries are referred to as floricane-fruiting or summer cropping cultivars. The annual fruiting cultivars also known as primocane-fruiting or autumn fruiting, autumn cropping, everbearing or tip-fruiting, complete the life cycle in one year (Heide and Sonsteby, 2011). A third group of cultivars, the so-called tip flowering ones, produces few flowers and consequently fruits at the end of the first growing season, but only in the tip of the top shoots while the remaining flower buds fruit the following (second) year. Depending on environmental conditions, biennial and annual cultivars can sometimes show the same behaviour and become tip flowering cultivars (Heide and Sonsteby, 2011).

Increasing popularity of autumn-fruiting raspberries, in which late season fruit is harvested from berries forming on the upper nodes of primocanes (Jennings and Brennan, 2002), has extended the production season and the period of attack of some foliar and cane pests. Some very early spring fruits with high value can also be obtained from the remaining lower nodes of these over-wintered primocane-fruiting types. Primocane-fruiting raspberries tend to be grown in the warmer areas of Europe where the temperature in autumn is relatively high and there is little risk of early autumn frosts. In primocane-

fruiting varieties, flowering and fruiting occur on distal buds with those proximal to the stem remaining dormant until next spring so that growers, especially in areas with little or no chilling hours, can have two crops per year, although characterized by lower quality fruit compared to an autumn crop only (Pritts, 2008). Primocane-fruiting cultivars present many advantages over traditional floricane-fruiting varieties as they do not have long chill requirements and so can be easily grown in warmer regions. They are easier to cultivate since at the end of the fruiting season, an entire plantation can be mowed to the ground, reducing pruning costs as well as those for pest and pathogen management since their life cycle is disrupted. Moreover, with the complete removal of the canes in late autumn, winter injuries are not an issue to the canes of primocane varieties. Floricane-fruiting cultivars require trellis to increase light penetration in the lower canopy to support their primocane (i.e. the non-fruiting vegetative canes formed during the first year of the biennial growth cycle) to fructo-canes (i.e. second year fruiting canes) development. For the primocane-fruiting varieties such structures, trellis, are not required because they only fruit at the distal end of canes (Pritts, 2008).

Primocane-fruiting cultivars cropping season can be extended by adopting different agronomic practices such as the use in spring of row covers, before canes emergence and until these reach the 0.5 m of height, to warm both soil and plants. Another practise to delay harvesting is abundant mulching after the canes have been removed. The mowing of the first new emerging primocanes in early spring and the tipping of the canes before flowering can delay harvesting time too (Pritts, 2008). Interest has also been shown in extended-season production under glass or under plastic structures in the northern European countries such as Belgium (Meesters and Pitsioudis, 1999) and the UK and now in the Mediterranean fringe (e.g. Spain and Greece). This trend will affect their pest and disease status. To satisfy these production systems, long primocanes grown in northern regions, such as Scotland, are lifted, chilled and stored for long periods before planting in late spring for late summer harvest under plastic tunnels production. The concept of extended-season production has been so successful that by careful manipulation of plant dormancy cycle and flower initiation it is now possible to produce fresh raspberries in Europe almost all year round.

1.2 Flower and fruit development in raspberry

Jennings (1988) states that a raspberry flower has five sepals and five petals with the latter being small and white even though in many *Rubus* species they are various shades of pink. Raspberry stamens arise in two crowded whorls in numbers ranging from 60 to 90, a smaller range than that shown by the styles (Figure 1.1). The numbers of both stamens and styles in the raspberry are affected by ploidy, genotype and major genes. Genes *M* and *F*, whose segregation can confer a dioecious habit, are very important. The recessive gene *f* suppresses the development of female parts and the recessive gene *m* suppresses the male parts. Most raspberries have genotype *FM* and are hermaphrodite, but *fM* genotypes are male, *Fm* are female and *fm* genotypes are neuter. The styles arise spirally on the terminal part of the receptacle, whose size and shape consequently determines the size and the shape of the fruit. Selection for larger fruit has given larger receptacles furnished with more styles: whereas in older cultivars the number of styles and hence the potential number of drupelets is rarely above 60.

While older cultivars presented flowers with number of styles, and hence of potential drupelets, rarely above 60; new varieties, selected for bearing larger fruit, have larger receptacles with higher number of styles.

Raspberry pollen grains differ in size and pore number depending on the ploidy with the polyploid species having bigger grains and a greater number of pores (maximum six). Flowers in raspberry secrete high amounts of nectar, in a diurnal rhythm up to 33 mg of nectar with 50% of sugars, making them very attractive to insect pollinators (Jennings, 1988). Raspberry flowers rewarded pollinators by producing abundant nectar that is freely exposed. In general, 10-12 µL per flower is produced and maintained in a liquid state even on warm days. For such features, these flowers are ideal for generalist pollinators such as honey bees (Nielsen et al., 2017).

Jennings (1988) reports that after the fertilisation, each ovary develops into a drupelet, which can be regarded as a miniature drupe. A drupe fruit is defined as one which develops entirely from a single ovary, and *Rubus* fruit are aggregates of drupelets formed by the ripening together of a number of ovaries all from the same flower and adhering to a common receptacle. Essentially, each drupelet is a complete fruit in itself and a miniature homolog of such drupe fruits as the cherry, plum or peach.

Once fertilized each ovary develops into a drupelet and raspberry fruits are composed of multiple fertilized ovaries from the same flower that aggregate to each other and adhere to a common receptacle. The drupelets are kept together by the entanglement of microscopic hairs, more specifically unicellular trichomes, which are particularly abundant on the side and base of the drupelets; such epidermal hairs are the only element guaranteeing drupelet cohesion (Jennings, 1988).

The fleshy part of the drupelets consists of parenchymatic cells with a thin cell wall that radiates from the pyrene, followed by a central region of oval-shaped cells and a thin layer of two/three collenchymatous hypodermis cells just below the epidermis. Soon after pollination, the fertilized ovaries undergo first a period of intense cell division (10-12 days) followed by a period of much slower growth in which the embryo develops while the endocarp hardens and finally the last stage is an intense cell enlargement causing a fast growth of the fruit. In general, all drupelets ripen at the same time, in about 30-36 days after pollination (Jennings, 1988)

Figure 1.1: Close-up pictures of respectively a raspberry fruit (left) and raspberry flower (right).

1.3 'Crumbly' fruit

A condition known as 'crumbly fruit' occurs to differing degrees in different raspberry varieties and is an indication of partial failure in the physiological processes of fruit development. 'Crumbly fruit' is linked with pollen abortion and embryo sac degeneration causing drupelets to be generally reduced in number but greatly enlarged in size, or, in the case of small reductions (Figure 1.2), cohere imperfectly so the fruit crumbles when is picked up (Graham et al., 2015).

Figure 1.2: Close-up picture of 'crumbly' fruit at two different developmental stages.
Pictures of two crumbly fruits at two different stages, green berry (left) and red berry (right). The fruits present few drupelets but in general with bigger size compared to normal fruits.

In reality 'crumbly' fruit is a term used to describe different conditions and so is too generic and consequently two new definitions were introduced at the beginning of this work with the aim of providing a standard definition of 'crumbly' fruit based on the different levels of severity at which the phenomenon presents itself. The Crumbly Fruit Condition (**CFC**) defines only plants where all fruit are symptomatic year after year and so plants are affected by an undefined disease or genetic disorder. Plants that display malformed fruits which look similar to 'crumbly' fruits in their appearance but where the symptoms are intermittent within a year or over the subsequent years were defined as plants with symptoms of Malformed Fruit Disorder (**MFD**).

Malformed Fruit Disorder is present when plants display uneven fruit set, generally at the very beginning of fruit production and this occurs mainly on the top lateral shoots. Fewer symptoms are observed as the season progresses and more flowers and fruits are produced by the plant resulting in little or no loss of yield. This condition can be more severe when

malformed fruits are observed throughout the entire fruiting season and displayed on most or all the laterals. However, these plants do not display the symptoms every year.

Malformed Fruit Disorder causes serious concerns for growers as it appears intermittently and is the type of crumbly fruit studied here.

1.3.1 Potential triggers of 'crumbly' fruit

There have been many causes suggested for the crumbly fruit phenomenon:

- failed pollination (i.e. too few ovaries fertilized)
- damaged flowers (i.e. over-visiting by pollinators)
- adverse environmental conditions during key stages of flower/fruit development (e.g. short day (SD) and low temperature (LT) favouring female sex expression)
- virus infections
- extensive tissue culturing
- genetic predisposition
- insufficient chill requirements
- phytoplasma
- excess flower nectar.

1.3.1.1 Viruses

Virus infections have been shown to be a cause of crumbly fruit in red raspberry. In their experiments, with a Meeker, Quito-Avela et al. (2014) tested Raspberry Bush Dwarf Virus (RBDV), Raspberry Leaf Mottle Virus (RLMV) and Raspberry Latent Virus (RpLV), alone or in all possible combinations, and they demonstrated how these viruses are responsible for very severe forms of crumbly fruit. However, not all plants affected by 'crumbly' fruit have virus infections.

1.3.1.2 Pollination

Pollination represents a decisive event in plant reproduction where pollen is transferred from the anthers to the stigmas of the carpels triggering the process of ovule fertilization that leads to the formation of seed and fruit. With anthesis, the flower opens and simultaneously anthers release pollen grains and when these reach the stigmas, they germinate due to a secretion, indispensable for their rehydration, produced by the papilla cells of the stigma. The pollen tube, emerging from the pollen grain, grows and elongates through the transmitting tract of the style to reach the ovule. Mature pollen carries two generative nuclei and when the pollen tube enters the micropyle (an opening on the surface of the ovary) it ruptures and releases the two generative nuclei so that while one unites with the egg cell to form the zygote, the other fuses together with the two polar nuclei of the embryo sac forming a triploid nucleus; the zygote cell produces the embryo while the triploid nucleus gives rise to the endosperm (Ramirez and Davenport, 2013). It seems evident that any process interfering with pollination cannot be excluded from the list of potential causes of 'crumbly fruit' since the lower the number of carpels fertilized the higher the chance of developing crumbly like misshapen fruits with a lower number of drupelets.

Raspberry presents as multi-carpel flowers, which means they have numerous pistils around a central core (i.e. receptacle). The more pistils that get fertilized the better for fruit development. Raspberries are completely self-fertile but pollen from the outermost whorl of anthers cannot reach the centremost pistils of the same flowers and so raspberry flowers still depend on insect pollination to produce normal shaped berries (Saez et al., 2014). To deal with the problem of poor fruit set, honeybee (*Apis mellifera*) and bumble bee (*Bombus terrestris*) hives are introduced into the field by growers. Field experiments have shown how honeybees and bumble bees are the main visitors to raspberry flowers and how they, especially honeybees, increase the deposition of pollen on the stigmas of carpels; the first step towards ovule fertilization. In their experiments, Saez et al. (2014) found a negative correlation between pollinators (honeybees and bumble bees) flower visit frequency and the number of drupelets set per fruit, as well as a negative correlation with the proportion of pistils damaged by pollinator visits on flowers and the number of drupelets set. Bumble bees proved to be the most damaging and where particularly abundant, they showed voracious behaviour to collecting nectar. This resulted in breakage of sepals of flower buds before anthesis, completely compromising the possibility of fruit

set and development. It is clear that any event causing damage to flowers, including pollinators over visiting, must be considered in regard to possible factors responsible for 'crumbly fruit' development.

The nectar standing crop represents the distribution and amount of nectar that a plant produces to attract insect foragers (e.g. bees) at a given time. It varies significantly in the plant and is influenced by the activity of insect foragers as well as by plant rate in producing nectar. Plants economize the production of nectar because it is energy demanding and because by modulating the distribution of it along the plant (i.e. combine flowers nectar-less and other filled with nectar) they favour cross pollination. In fact, when pollinators encounter individual flowers or plants free or low in nectar, they quickly move to other distant patches of flowers, belonging to other plants. This indirectly reduces the probability of geitonogamy (i.e. self-pollination between flowers of the same plant) and consequent risk of inbreeding (Nepi et al., 2018).

Moreover plants can modify nectar composition to attract pollinators and do so by varying both primary metabolites (i.e. sugars and amino acids) and secondary metabolites (i.e. nicotine and caffeine) content (Nepi et al., 2018). In a situation of lack of pollinators, such as those that might be found in glasshouses and polytunnel, plants may increase the production of nectar as well as the content of sugars (Gardener and Gillman, 2002), to try to attract pollinators. Such nectar is then produced in excess since the lack of foragers prevents its removal from the flowers and so it drips over the carpels preventing pollination. The result can be potentially cause crumbly-like, misshapen fruits with reduced number of drupelets. The excess of nectar can even cause fungal proliferation leading to diseases such as sooty mould.

Some cultivars appear to be more prone to the condition than others but in cases where the environmental conditions differed from the normal seasonal level, random symptoms of crumbliness could be displayed in cultivars not previously known for the problem.

1.3.1.3 Flower development

In general, in both monoecious and dioecious plants, the expression of female flowers is favoured by short day (SD) and low temperature (LT) conditions with these environmental cues causing changes in endogenous hormones, primarily auxin and

gibberellins, with auxins promoting femaleness and gibberellins male expression. An increased femaleness in raspberry flowers might be triggered by SD and LT that are responsible for reducing the concentration of gibberellins in the shoots and, as stated previously, this class of hormones is responsible for promoting shoot growth as well as acting as a male growth regulator (Woznicki et al., 2016). The same authors designed experiments with 'Glen Ample' canes grown under polytunnel. These canes at the end of the summer were moved to a phytotron where the plants were subjected to three different temperature treatments (9, 15 and 21 °C) for six weeks to induce floral initiation and differentiation. The following summer, the berries harvested were significantly different, in terms of number of drupelets per berry, between the three different treatments. The highest number of drupelets was found in the treatment with lowest temperature as a consequence of increased number of carpels in flowers initiated and differentiated with the 9 °C treatment (Woznicki et al., 2016). Exposing plants to higher temperatures at the end of the fruiting season during flower bud induction can potentially reduce the number of carpels differentiated and increase the chances of producing, the following fruiting season, misshapen crumbly like fruits with lower drupelets number; climate change and temperature rises may therefore be a further threat for raspberry growers.

These environmental effects are responsible, at a molecular level, for changes in the endogenous plant hormones levels with auxin favouring female and gibberellins male sex expression. Increased femaleness in the form of a boosting of carpels (pistils) per flower is a characteristic also in *Rubus* species with perfect (hermaphrodite) flowers. Crumbly condition could be the result of plants having disproportionate numbers of pistils/stamens. It seems possible that a very low number of pistils per flower (decreased femaleness), paired (or not) with a high number of stamens, could be the simplest scenario responsible for the crumbly condition. In fact, as long as too few ovaries get fertilized, the developing fruit can be misshapen and then unmarketable with consequent loss of yield for growers. Together with hormones, other molecules could participate in crumbly condition promotion, and within these there are ROS (Reactive Oxygen Species), mainly superoxide radicals (O_2^-) and hydrogen peroxide (H_2O_2), both are sub-products of cell respiration and are even plant stress response signals (Huan et al., 2016).

1.3.1.4 Genetic disorders in propagation

Genetic disorders affecting fruit production cannot be excluded as latent triggers of 'crumbly fruit' in red raspberry. Vegetative propagation can cause somatic mutations whose symptoms can occur after many cycles of multiplication and before they are actually discovered, resulting in propagation of material with genetic disorders, with 'crumbly fruit' being the most abundant of those causing reduction in raspberry plants fertility (Jennings 1988). Micropropagation of red raspberries, the application of tissue culture *in vitro*, presents several advantages over the traditional vegetative propagation methods, such as year-round production, pathogen-free progenies, large scale production and multiplication of recalcitrant species, and has been widely adopted. All the micropropagation protocols based on axillary branching should guarantee the production of true-to-type plants with genetic uniformity (Vujovic et al., 2017) but cannot exclude somatic mutations. Micropropagation results in a higher incidence of somatic mutations as the number of multiplication cycles is much higher, compared to conventional vegetative propagation techniques and so micropropagation must be considered as a potential trigger of 'crumbly' fruit.

1.3.1.5 Environmental factors

In biennial raspberry cultivars three environmental conditions are indispensable to allow the formation of adequate plant architecture suitable for high quality fruit production. These factors are, relatively high temperature in summer to promote ample shoot growth, low temperature at the end of the fruiting season to stimulate adequate floral bud initiation as well as dormancy induction and finally exposure to chilling to break dormancy. Exposure to chilling temperature in the range 0-7 °C, rather than subfreezing, for 6 - 8 weeks proves to be enough to beak dormancy of the uppermost buds while for those on the lower part of the cane, those producing longer fruiting laterals, require longer chilling periods. In annual fruiting varieties, where chilling requirements are lesser, exposing canes to such conditions promote flowering (Heide and Sonsteby, 2011). Climate change causing higher temperatures and a failure to deliver the chilling requirements might be responsible for inadequate dormancy breakage whose potential role in compromising fruit quality and in particular fruit shape cannot be excluded.

1.3.1.6 Phytoplasmas

Another 'crumbly fruit' potential trigger are phytoplasmas, plant bacteria with wall-less cells belonging to Mollicutes class. Phytoplasmas are pathogens that colonize up to 700 plant species (Hoshi et al., 2007). They are transferred by a range of processes including: phloem-feeding insect vectors, vegetative propagation (grafting or cutting) of infected plants, parasitic plant species (*Cuscuta* sp.), seed transmission and root anastomosis (Weintraub and Beanland, 2006). In raspberry, but in general in the whole genus *Rubus*, phytoplasmas are responsible for a disease called *Rubus* stunt whose symptoms are stunting, witches broom, small leaves, short internodes, enlarged sepals, phyllody (abnormal development of floral parts), flower proliferation and malformed fruits (Linck and Reinecke, 2019). In *Rubus* stunt, the proposed agent is '*Candidatus* Phytoplasma *rubi*' and its vector is the leafhopper (*Cicadellidae*) *Macropsis fuscula,* a phloem-feeding insect of the order hemiptera that can be found in Western Europe, Ukraine, European and far east Russia, Caucasus, Kazakhstan, middle Asia, Japan, Western Canada and USA. The insect acquires the bacterium during feeding but cannot transmit it before the phytoplasma transfers from the gut to the salivary glands Plants already affected by other pathogens, especially raspberry mosaic complex (i.e. a disease caused by different viruses), when infected with phytoplasmas can display symptoms of *Rubus* stunt that are even more severe causing death within a year from the infection (Linck and Reineke, 2019). As stated above, one of the symptoms associated with *Rubus* stunt is the formation of malformed fruits suggesting that phytoplasmas cannot be excluded as another potential trigger of 'crumbly' fruit in red raspberry.

1.4 The molecular biology behind flower and fruit development

Flowering is regulated by a range of environmental and physiological signals (Fornara et al., 2010, Song et al., 2012) that still need to be fully understood in perennial crops. CONSTANS (CO), a key component in leaves of the photoperiodic pathway, is stabilised by light and accumulates under long day conditions. CO activates transcription of FLOWERING LOCUS T (FT) (Simon et al., 2015) which interacts with bZIP transcription factors (Abe et al., 2005, Cao et al., 2016a) activating a cascade of downstream genes leading to flowering. This basic flowering process is impacted by several autonomous and stress related signals. For example, the MADS box FLOWERING LOCUS C (FLC) and short vegetative phase proteins (SVP) form a complex to repress flowering until the plant is exposed to the appropriate level of cold. In raspberry, RiMADS_01 was identified as a potential candidate affecting vernalisation (Graham et al., 2009).

The genetic basis of raspberry fruit development is still not well understood, although some studies have been carried out to look at overall control of fruit development and aspects of ripening (Paterson et al., 2013, Dobson et al., 2012, McCallum et al., 2010, Graham et al., 2009, Simpson et al., 2017, McCallum et al., 2013). Recently some progress has been made in understanding 'crumbly fruit' by determining that both genetic and environmental factors affect the crumbly phenotype. A location on Linkage Group 1 (LG1) and another on LG3, highly significant for determination of the 'crumbly' fruit syndrome in red raspberry, have been identified and they are robust across seasons and in different environments (Graham et al., 2015).

The current model of regular fruit set implies that ovary growth is blocked before pollination and that auxin is a key regulator of ovary growth de-repression at fruit set. Other phytohormones such as gibberellin, cytokinin, brassinosteroids, ethylene and abscisic acid have been identified as having roles in fruit initiation and development; (Goetz et al., 2007, Pandolfini et al., 2007).

Fleshy fruits are botanically diverse in the way they develop, whereas raspberry, tomato and grape are derived from the ovary, other fruits such as apple and strawberry are derived from the receptacle or from the expansion of the sepals. Despite these botanical differences, all fruits undergo similar developmental steps including fruit set, growth, maturation and ripening. Fruit set is the first stage of the development after the

fertilisation event and is followed by an active cell division and expansion phase, called growth, that causes the fruit to attain its maximum size. This is followed by a stage where fruit acquire the prerequisite competence to enter in the final developmental stage (i.e. ripening). Fleshy fruits are physiologically classified as climacteric and non-climacteric with the first showing concomitant increase of respiration and ethylene biosynthesis upon initiation of ripening while non-climacteric fruit lack these two attributes at the onset of ripening (Kumar et al., 2014) .

The combined action of three hormones: auxins, gibberellins (GAs) and cytokinins play a major role in the regulation of fruit set. Importantly, auxin, GA and cytokinin levels increase at fruit set and the requirement of their higher levels has been already validated in tomato by exogenous treatment, which causes parthenocarpic fruit formation (Kumar et al., 2014). These authors suggest that auxins promote fruit set and growth, at least partially, by controlling the GA levels. Differential regulation of many genes related to biosynthesis and signalling of other phytohormones such as, ethylene, ABA, cytokinins and brassinosteroids (BRs) further suggests that besides auxins and GAs, these hormones are also important during fruit set and early fruit development stages (Ruan et al., 2012).

From a molecular point of view, fruit set involves the concerted action of auxins and/or GAs and/or cytokinins or BRs through the biosynthesis and/or the signalling which regulates the activation of core cell cycle genes during early fruit development stages (Goetz et al. 2007; Pandolfini et al. 2007). ABA and ethylene play antagonistic roles in fruit set even though such mechanisms remain unidentified (Kumar et al. 2014).

Seed and fruit development are intimately connected and synchronized. Seeds are rich sources of hormones, auxins, GA and cytokinin all of which are involved in stimulating the growth of the surrounding tissues and also in determining the fruit size. Interplay between these hormones is necessary for fruit growth and, so far, the established role of auxin, in the regulation of cell expansion, seems to be the most important during this development phase. Once cell division stage is over, auxins and GAs become the main regulators of the cell expansion phase. These hormones are localized mostly in the seeds and then transported to the surrounding tissues but, except for the auxins, the knowledge on this aspect remains limited (Kumar et al., 2014, Sablowski and Dornelas, 2014).

Pomares-Viciana et al. (2017), demonstrated the role of auxin in regulating fruit development after pollination. The authors, studying the parthenocarpic development process in zucchini, found fruit development to be closely related to the activation of

auxin response genes such as: ARFs (Auxin Response Factors), Aux/IAAs and TIRs (TOLL/Interleukin-1 receptors) encoded by multigene families. Transcriptome analyses of *CpARF8, CpIAA9* and *CpTIR1* have revealed that they show tissue-specific expression, maintaining the structure and function showed in other species. Such key auxin signalling genes showed a specific level of quantified mRNA in pre-anthesis and anthesis that changed after the fertilisation cue, supporting their role in the preparation of the ovary to become fruit in zucchini. A similar scenario could be found in raspberry where, among other things, the potential development of parthenocarpic drupelets cannot be completely excluded. A similar scenario, involving auxins in the regulation of fruit growth and development after pollination, could be hypothesised in raspberry too. In fact, the complex structure of its berry, containing many individual drupelets held together, could require even parthenocarpic growth, of part of the ovaries composing the fruit, simplifying, an otherwise complicated fruit development process depending upon too many fertilization occurring together at the same time.

1.5 Raspberry plant health certification scheme

Raspberry plants can be infected by many different pathogens and the eventual emergence of diseases depends on many factors. Raspberry is a perennial crop requiring initial high investment costs in plantation establishment and maintenance thus the use of pathogen free plants is important for productive and profitable plantations. Over a number of seasons plants grown in field conditions become increasingly infected with pests and pathogens with negative effects on yield. It is very important that growers maximize plantation lifespan and yield, and reduce crop losses by re-stocking their plantations or establish new plantations with certified, pathogen-free planting material (Dolan, 2018).

The Plant Health Certification Scheme (PHCS) was introduced in UK in the 1940s and since then it became a significant contributor to the success and development of the raspberry industry. The PHCS provides growers and propagators with materials derived from plants tested for varietal identity, health status and vigour. The health of the plant stocks is maintained by introducing pathogen tested plant material into the propagation system, keeping a constant history of the certifications and by applying restrictions for eligibility in each grade of the certification scheme. In 2017 the UK Plant Health Certification Scheme was harmonised with the EU scheme creating a single regulation for all the countries that legislates about registration of suppliers and varieties, labelling for each grade and certification and gives the rules for official inspections. The EU scheme defines five grades: Pre-basic, Basic 1, Basic 2, Standard and *Conformitas* Agraria *Communitas* (CAC) – Plant Passported which replaced the corresponding previous UK certification grades: Nuclear Stock, Super Elite, Elite, Certified and Non-Certified (Dolan, 2018).

The James Hutton Institute is the only UK provider of buds and roots for entering the plant health certification scheme from pest- and pathogen-tested *Rubus* Pre Basic grade ‘mother’ plants. Mother plants of established varieties or newly released ones enter the scheme after being tested (i.e. biological and/or molecular) for a range of different diseases. These plants are maintained in sterile compost inside an insect-proof glasshouse, irrigated with UV sterilised water and are visually inspected and tested; if they fail one test they are destroyed. The aims of the certification scheme are two fold, to guarantee, to both growers and propagators, high quality planting material in terms of health, vigour and purity, while the second goal is to avoid the spread of harmful pests and diseases by introducing only pathogen tested material with an unbroken history of certification and

limiting the time for stock material to remain eligible for certification. The 'crumbly fruit' condition is assessed as part of the plant health certification scheme. The aim is to release only material from mother plants that bear only uniform and stable fruits. Plants are grown in optimum conditions and pollination efficiency is achieved by use of commercial bee-hives. When plants fruit, they are visually inspected and all those displaying atypical phenotypic characteristics or producing crumbly or misshapen fruits are destroyed so no further propagation is allowed (Dolan, 2018). Given the occurrence of the condition in different seasons this may only be partially successful at eliminating the condition as only symptomatic plants will be destroyed.

Planting of certified *Rubus* stock material disease-free in clean soil, where persistent fungi, bacteria, virus and certain pest have been removed, represents the best strategy to increase plantation lifespan. Important crop management strategy is the removal (i.e. mechanically or chemically) of suckers and spawn in the inter-row space to prevent overgrowth in rows as well as to avoid competition for light, water and nutrients; the number of primocanes for the following fruiting season is controlled by pruning in winter and early spring to reduce inter-cane competition. The canes are grown on a post and wire system that helps to carry the weight of the fruits and to protect crops from wind, harvesting and general cultivation practices. Soon after the end of the harvest, the old fruiting canes are removed to eliminate from the plantation potential sources of fungal inoculum; a fundamental practice to keep longer the health of the crops (Graham and Brennan, 2018a).

1.6 Raspberry breeding

In raspberry, the modern cultivars, although they remain only few generations away from their wild progenitors are, as a consequence of domestication, much more homozygous and genetically similar to each other. This lack of diversity can be a threat when it comes to seeking specific traits, for example those related to resistance to pathogens and pests, for breeding programs. The introduction of unselected raspberry clones and species is a strategy that can be used by breeders to increase the genetic diversity and help facing the future crop challenges primarily related to climate change and emergence of new pests and diseases (Graham and Brennan, 2018a). Collections of *Rubus* accessions both wild and cultivated gene-pools are an important source of genetic variability that can be used to implement breeding programs; the most important collections are found in North America (i.e. National Clonal Germoplasm Repository of the United States Department of Agriculture and the Canadian Clonal Genebank) and Europe (i.e. The James Hutton Institute and the East Malling Research, respectively located in Scotland and England).

The raspberry industry currently relies on only a limited use of chemicals. No high quality varieties that combine pest and pathogens resistance are available so the integration of new traits adding resistance/tolerance to biotic stresses is becoming a priority for all breeding programmes. In the selection of new varieties, the new cultivation practices (e.g. protected crop systems and production on pots and container) are obviously highly regarded. Emerging technologies such as new phenotyping platforms that use imaging for detecting phenotyping variation are becoming available for breeders (Jennings, 2018).

Due to retailer demands mainly but also consumers preferences, the raspberry industry relies on a small number of cultivars that among other things have limited resistance to both pests and pathogens (Finn and Hancock, 2008). The increasing lack of availability of many pesticides is causing serious concerns for crop management. The increased awareness of consumers on topics like ecological footprint of agricultural practices, especially those directly linked with food production, is also pushing the industry towards the adoption of more sustainable agricultural practices. Breeding represents a tractable and durable solution to this problem since the selection of new varieties with enhanced traits can allow growers to meet the markets demands. Conventional breeding is too slow to keep pace with the new challenges set by global markets, especially for highly heterozygous outbred such as raspberry cultivars (Finn and Hancock, 2008).

The main challenge related to breeding is the time-consuming nature of the whole process that foresees the collection of quality and yield data on seedling population. Accurate measurement of yield data requires hand picking which is costly and cannot assess potential machine harvestability. An alternative could be the visual scoring of yield, but it can reduce genetic improvement and compromise cultivar development. Recent works have highlighted how berry weight and shoots lateral length, when measured in the first fruiting season, can allow accurate prediction of the following years and so can be used by breeders to focus on genotypes with such features when they select material for breeding. A positive correlation was found between yield and the combination of cane length plus the number of broken buds per cane in primocane varieties (Jennings, 2018).

Fruit quality is another important priority of any breeding programme with flavour, colour brightness, size and shape as the prime attributes influencing consumers acceptance: to ensure repurchasing, fruits must be pleasing in terms of appearance and taste. Growers are also interested in fruit of consistent big size since they contribute to a reduction in harvesting costs. Prolonged shelf-life in punnet for at least seven days after harvesting is a preferred trait for retailers (Jennings, 2018).

Raspberry production destined for fruit processing, would benefit from cultivars adapted to machine harvesting as this would lower labour costs. Machine harvesting varieties must present special characteristics, spine-free strong, long canes with fruit well-presented and exposed to the outside of the bush are required to guarantee repeated passes of the harvester across the season while minimising the damage to the following years canes. High quality fruit that are easily shaken from the receptacle during harvesting is a priority and obviously for fruits destined for Individual Quick Frozen (IQF) it is important that the berry maintains its integrity while processed on the freezing line. In floricane-fruiting varieties, machine harvestability seems to be associated with other traits such as large berry weight and high number of shoots laterals for early-ripening cultivars while number of shoots laterals and long cane length for late ripening varieties; such aspects must be considered when selecting new floricane-fruiting cultivars (Jennings, 2018).

The two main breeding programmes in UK are led by the East Malling Research (EMR) and the The James Hutton Institute (JHI). EMR, based in south England, is responsible for the ‘Malling’ series of raspberry. ‘Malling Promise’, ‘Malling Exploit’ and, in particular, ‘Malling Jewel’ were the first varieties produced by this research centre. In 2008 EMR introduced ‘Octavia’ a late floricane-fruiting that helped to extend the season, in UK and particularly in Scotland, by closing the gap between summer and autumn fruit

productions. More recently, East Malling Research recently concentrated on developing primocane cultivars including 'Autumn Treasure' with good quality fruit and some tolerance to *Phytophthora rubi* and more recently with 'Malling Bella' and 'Malling Charm' (Jennings, 2018). The James Hutton Institute is responsible for the 'Glen' series, the first variety released was 1969 was 'Glen Clova' after that, early in the 80s, two new cultivars were produced, 'Glen Moy' and 'Glen Prosen' until in 1996 when it released the most iconic and successful cultivar 'Glen Ample'; for many years the standard in the European wholesale for both fresh and processing markets. More recent varieties of the 'Glen' series are 'Glen Fyne' with good machine harvestability features, 'Glen Ericht' with good tolerance to *Phytophthora ruby* and 'Glen Dee' and 'Glen Carron' offering large fruit with long shelf-life. Recently 'Glen Mor' has been released with resistance to root rot using marker assisted selection. In the last few years, the JHI programme has also focused on early autumn fruiting primocanes with increased picking efficiency to reduce labour costs and extend the season. In continental Europe, Poland is one of biggest raspberry producers and NIWA berry breeding Ltd released important varieties, in particular 'Polana', 'Polska' and more recently 'Polonez', 'Poemat' and 'Delniwa' while 'Laszka', 'Radziejowa' 'Sokolica' and 'Przehyba' (Jennings, 2018).

The raspberry industry has been revolutionized by primocane varieties since their annual life cycle allows production in areas where extremely cold winters could damage the canes of floricane-fruiting plants or where insufficient chill would compromise the production of summer bearing cultivars. Based on primocane varieties, a new cultivation system emerged in California where primocane plants are in production only 18 months because this was proven to increase fruit quality and its shelf life. Such cultivation system is made possible by the peculiar, one and a half year, growth cycle of primocane cultivars with the canes top laterals fruiting early in Autumn the first year while those on the bottom early in summer the following year (Jennings, 1988). This new management system boosted the Californian raspberry industry and was then replicated in warm temperate areas all around the world: South and Central America, Southern Europe, Australia and South Africa.

Recently in terms of breeding primocane varieties, companies have launched programmes with restricted licensing availability and exclusivity. This has resulted in the release of many new varieties. 'T-plus' and 'Diamond Jubilee' in UK, 'Sapphire' 'Pearl' and 'Jade' by a British-American partnership. In the Netherlands: 'Imara', 'Kweli', 'Kwanza', 'Mapema' and 'Rafiki'. In France, 'Paris' and the very large fruit bearer 'Versaille' while

Spain produced two cultivars, ‘Lupita’ and ‘Adelita’ both grown successfully in Morocco. In Italy, the most successful cultivars released are ‘Castion’, ‘Enrosidira’, ‘Vajolet’ and ‘Lagorai’ (Jennings, 2018).

1.7 The use of genomic technologies in raspberry research and breeding

Molecular breeding techniques have advanced research in raspberries. Molecular markers are now used to detect genome-wide variability in coding and non-coding regions and produce linkage maps that containing numerous genetic markers useful in marker assisted breeding (Gupta et al. 1999; Koebner and Summers, 2003; Morgante and Salamini, 2003) (McCallum et al., 2018).

Genetic markers are genetic differences (polymorphisms) between individuals of the same species or between individuals of different species. Genetic markers, in general, do not represent the gene target, they are simply tightly linked with it and in fact they can be referred to as gene tags. Genetic markers generally do not affect the phenotype of the target trait and they occupy specific genomic regions (loci) along the chromosomes (Collard et al., 2005). Three different genetic markers can be distinguished: morphological markers (also called classical or visible markers), biochemical markers and DNA or molecular markers. The morphological markers are a phenotypic character that can be easily visualized (i.e. petals colour, seed shape, etc.) while the biochemical markers primarily are differences in enzymes that can be detected by electrophoresis separation or and specific staining. The DNA markers are sites along the genome where variations in the DNA sequence (polymorphism) occurs (Collard et al. 2005). The main problem with morphological and biochemical markers is that they are limited in number and are highly influenced by environmental factors or specific developmental stages, while DNA markers, due to their abundance, are by far the most used markers. They can be caused by different types of DNA mutations; substitution mutations (point mutations), rearrangements (i.e. insertions or deletions) or error in the replication of DNA sequences resulting in regions repeated in tandem (Paterson 1996b). The DNA markers are selectively neutral because they are mainly located in non-coding regions of the genome, they are almost unlimited in number and they are not affected by environment or development stage. DNA markers are not just used for constructing linkage maps, but they have found important application in assessing the level of genetic diversity and in cultivar identification and so are widely used selecting parents in breeding (Collard et al. 2005). According to the method utilized for their detection, DNA markers can be divided further in three classes, hybridization-based, polymerase chain reaction (PCR)-based and DNA sequence-based. Of particular interest are those DNA markers that allow discrimination of individuals of the same or different species; these markers are called

polymorphic markers while those that cannot distinguish between different genotypes are called monomorphic markers. Polymorphic markers can be further divided in codominant and dominant on the basis of their ability to distinguish between homozygotes and heterozygotes; codominant markers may have many different alleles while dominant markers only two alleles (Collar et al. 2005).

Many different markers have been developed for the construction of linkage mapping and the most common are, amplified fragment length polymorphism (AFLP), random amplified polymorphic DNA (RAPD), simple sequence repeat (SSR) and single nucleotide polymorphisms (SNPs) (Antonius-Klemola, 1999; Hokanson, 2001; Graham et al., 2002; Stafne et al., 2005; Woodhead et al., 2008; Castillo et al., 2010; Dossett et al., 2012; Bassil et al., 2014). Linkage maps with markers strongly associated to gene(s) or quantitative trait loci (QTL), linked with specific traits of interest, are used for both breeding and to study the genetic controls of specific phenotypes of interest. The advent of genotyping by sequencing (GbS) has advanced linkage mapping with the creation of high resolution and marker-saturated maps (McCallum et al., 2018). In red raspberry two important GbS maps have been developed, one from a mapping population between 'Heritage' and 'Tulameen' and one between Latham and Glen Moy ; these new maps allowed the confirmation and greater resolution of established QTLs and also identified new ones (Hackett et al., 2018).

Linkage maps have already been widely established as an important tool to facilitate both breeding programs and genetic studies since they allowed the linking of traits of interest to markers or in some cases even to genes responsible for trait variations to chromosomal loci (McCallum et al., 2018). The challenges that climate change poses together with demand for more sustainable agricultural productions are all pushing research and breeding towards the selection of new varieties with improved adaptability to the new climatic stress and with increased resistance to biotic stresses (McCallum et al., 2018).

In red raspberry, there are both natural and experimental populations that show adaptation to different habitats and variability above many and different traits. The implementation of gene maps, with increased marker saturation, together with the availability of further new genome sequences, opens the doors to genome wide association studies (GWAS) and the analysis of more diverse populations with potential increases in allelic diversity and improved resolution; provided that genetic heterogeneity and marker density, in raspberry, is sufficient to identify genes or loci by association. Together with genomic approaches, transcriptomic and metabolomic methods can also be utilized to support

linkage and/or association mapping (Gupta et al., 1999; Morgante and Salamini, 2003) (McCallum et al., 2018).

1.8 Aim of the thesis

'Crumbly' fruit is a very generic term used to describe raspberry plants bearing uneven fruit with reduced number of drupelets, which in general tend to be larger in size and to not adhere perfectly to each other so when the fruit is picked it crumbles, hence the name 'crumbly' fruit. The way the plants can display 'crumbly' fruit symptoms can vary. In order to focus this work, two new definitions were established. As some plants always display 'crumbly' fruit, and this occurs every year, those plants will be classified according to the new definitions introduced here as being affected by Crumbly Fruit Condition (**CFC**); the genetic related form of 'crumbly' fruit. Other plants bear uneven fruits only at the beginning of the fruiting season and, in general, only on the top laterals or, in case the symptoms are exhibited along the whole fruiting season and on all the laterals, however it never happens that all the fruits are 'crumbly' and it does not occur every year. Such plants, according to the new definitions are affected by Malformed Fruit Disorder (**MFD**) a more environmental related form of 'crumbly' fruit (details about the new 'crumbly' fruit definitions can be found in section 1.3 of this chapter). The MFD condition is the subject of this thesis with the seasonal variation and thus impact on the industry being the main concern of raspberry growers. Despite that and only for simplicity, from here onwards, the generic term 'crumbly' fruit will be used through the whole document unless otherwise stated.

Many potential causes have been associated to 'crumbly' fruit in red raspberry (details about these 'crumbly' fruit potential triggers listed below can be found on section 1.3 of this chapter):

- failed pollination (i.e. too few ovaries fertilized)
- damaged flowers (i.e. over-visiting by pollinators)
- adverse environmental conditions during key stages of flower/fruit development (e.g. short day (SD) and low temperature (LT) favouring female sex expression)
- virus infections
- extensive tissue culturing of plants may increase the emergence of the condition
- genetic predisposition
- insufficient chilly requirements
- phytoplasma
- excess of flower nectar.

Jennings (1967b) stated that 'crumbly' fruit is an indication of a partial failure in one or more physiological processes concerned with fruit development. It is clear that a better understanding of the physiology and molecular processes behind fruit development is required to help develop a strategy for controlling, or ideally eradicating, 'crumbly' fruit and this forms the subject of this thesis.

One hypothesis in this work concerns a hormonal coordination process regulating and synchronizing the growth of all the fertilized ovaries. Without this system the first fertilized ovaries would start to grow earlier, creating a gap with those fertilized later that will be never filled up, resulting in fruits misshapen and/or crumbly. The centre of this regulatory process is postulated to be the receptacle, that acting as a hub by means of molecular signals (e.g. plant hormones), establishes an intense hormonal crosstalk with the fertilized ovules that helps to coordinate the fruit development. This allows the last ovules to be fertilised to start growing simultaneously with those that received the pollen earlier. In simple terms, the receptacle acts like a switchboard operator taking note of all the calls it receives immediately after an ovary gets fertilised. In this manner, soon after the receptacle has received a number of "calls/messages" equalling the minimum number required to develop a proper fruit, it contacts all the fertilised ovules at exactly the same time and with its signal triggering their simultaneous growth. No published scientific papers are available to support this hypothesis and a range of experiments have been specifically designed to address this.

1.8.1 Outline of the study

Two different strategies were employed during this work,

1. The first focused on studying the differences in gene expression between 'crumbly' and normal fruit in progeny of a mapping population (Glen Moy x Latham) known to segregate for the 'crumbly' fruit condition across different seasons (Graham et al., 2015).
2. The second approach was to study the hormone regulation behind fruit development in artificially induced 'crumbly' fruits.

The first goal of this study was the identification of the genetic control of the condition through QTL mapping on the GbS map and examining the differences in gene expression

levels between normal and 'crumbly' fruit. To achieve this a linkage mapping and transcriptomic analysis was conducted on samples from progeny of a mapping population between two cultivars, Glen Moy and Latham which were known to segregate for the condition. It was proposed that identification of QTL and genes strongly associated with the 'crumbly' fruit phenotype together with the hormone profile work would assist in understanding of the control of the condition and to help develop a control strategy either through markers for marker assisted breeding and/or for the development of a robust test for the assessment of the likelihood of 'crumbly' fruit development to be used in the Plant Health Certification Scheme.

The work on the second strategy aimed to develop a method of inducing the 'crumbly' condition due to the unpredictable nature of the expression of the trait across seasons, to allow the study of hormone levels between normal and 'crumbly' fruit. This involved interfering with the regular growth of the fruit to induce the 'crumbly fruit' condition producing artificial misshapen crumbly like berries, thus allowing a controlled study of the condition. By considering pollination and fertilization as the initial steps of fruit development, two potential ways could be considered for interfering with the proposed, receptacle coordinating, fruit growth regulating process. Up-stream of pollination where, by physically damaging/removing the style and stigmas of the majority of carpels to a low number such that even though all the remaining intact carpels were pollinated, there would not be enough to allow the formation of a proper fruit, but rather one resembling a crumbly fruit. The other approach to prevent the proposed fruit regulating process mediated by the receptacle would be downstream of the pollination/fertilization of the carpels. In this case all the flowers must be pollinated to allow the maximum number of ovules to get fertilized, but mechanical damage to the receptacle where the ovules are attached would affect its regulating function compromising the normal growth of the fruit. In this work the second approach was used to test the validity of the proposed regulating mechanism of fruit development because addressing directly the receptacle would have been more congenial to investigate its potential role as grow mediator.

Hormones are key players in normal development of fruits (Kumar et al., 2014). In general, auxins, gibberellins and cytokinins, interacting in concert or alone, play crucial roles in the first two phases of fruit development namely, fruit set and fruit growth. These are the stages when the crumbly condition appears. Understanding the hormones involved in fruit set together with those activated during stress conditions, in crumbly and non-crumbly fruit is a key factor to better understand the trigger(s) of such condition and

possibly find solutions to monitor and then mitigate it. Therefore, a method of generating crumbly fruit on demand was required for hormone research as the resulting phenotype is the same as that produced commercially depending on season.

Chapter 2: Materials and methods

2.1 Plant material

A mapping population consisting of a full sib family generated from a cross between the European red raspberry cv. Glen Moy and the North American red raspberry cv. Latham; with the latter being the 'crumbly' parent. This population was planted in autumn 2001 in a randomised block design with three replicates and two plant plots.

2.1.1 Plant material 'crumbly' induction experiment

Long canes cv. Glen Ample were purchased from EU plants Ltd. (Abingdon, UK). Plants were cultured in controlled environment, plant growth room, (Nijssen, The Netherlands). For the first two weeks the following parameters were set: minimum temperature 10 °C, maximum temperature 14 °C, relative humidity 70% and daylight length 16 hours (from 7:00 AM to 11:00 PM); such conditions helped the plants to acclimatize. After the two weeks of acclimatization, all the environmental parameters remained unchanged except the maximum temperature that was increased by two degrees, from 16 to 18 °C.

2.1.2 Plant material 'crumbly' markers Genome Wide Association Study (GWAS)

Sixty-three different genotypes (i.e. raspberry cultivars and selections), forming part of the germplasm available at The James Hutton Limited raspberry breeding programme were selected. The full list of genotypes can be found in Appendix (see Table A.1).

2.1.3 Harvesting plant material

Plant tissues (i.e. closed bud, open flower and green berry) from two different phenotypes (i.e. mostly and never 'crumbly'), with mostly 'crumbly' indicating progeny assessed as being 'crumbly' 75% or more of the times scored (Graham et al., 2015). These plant tissues were harvested into 2 mL Eppendorf safe-lock tubes (Eppendorf AG, Germany) and flash frozen using liquid nitrogen. Harvested tissues were stored at -80 °C until required for molecular analysis.

2.1.4 Harvesting plant material 'crumbly' induction experiments.

Plant materials (i.e. receptacle and drupelets) of raspberry fruit at two different developmental stages (i.e. green and red fruit) and from two different types (i.e. artificially induced 'crumbly' and control) were harvested into 85 mm x 75 mm (height x width) Bryson Packaging™ Minigrip™ polyethylene write-on bags (Fisher-Scientific Ltd, UK) and then flash frozen using liquid nitrogen. Harvested tissues were stored at -80 °C until required for molecular analysis.

2.2 Molecular protocols

2.2.1 Nucleic acid extraction (isolation of RNA)

RNA was isolated from three different parts (i.e. closed bud, open flower and green berry) of full sib family generated from a cross between the European red raspberry cv. Glen Moy and the North American red raspberry cv. Latham. Minimum 70 to maximum 90 mg of plant material were weighed and then fine ground with mortar and pestle using liquid nitrogen. The extraction/purification was performed using the RNeasy Plant Mini Kit (Qiagen, Germany) and with the RNAse free and DNAse I Set (Qiagen, Germany), following the manufacturer's instructions.

2.2.2 Nucleic acid extraction (isolation of DNA)

Genomic DNA was isolated from buds of 63 different genotypes (see Table A1 in appendix for the full list) of raspberry (*Rubus idaeus*). Minimum 70 to maximum 90 mg of plant material were weighed and then fine ground with mortar and pestle using liquid nitrogen. The extraction/purification was performed using the DNeasy Plant Mini Kit (Qiagen, Germany), following the manufacturer instructions

2.3 Analysis of nucleic acids

2.3.1 Quantification of nucleic acids by spectrophotometry

Concentrations of DNA and RNA in 1 µL of sample were estimated using a NanoDrop® ND-1000 full-spectrum UV/Vis spectrophotometer (NanoDrop®, USA). The NanoDrop® uses a modified Beer-Lambert equation to correlate the calculated absorbance with the concentrations of the samples. DNA and RNA samples were measured at 260 and 280 nm with ratios of 1.8 and 2.0 being accepted respectively for DNA and RNA.

2.3.2 Gel electrophoresis

Agarose gels were prepared by mixing 0.75 g of agarose with 50 mL of 1X Tris-borate/EDTA (TBE) buffer. The mixture was heated in a microwave on medium power for 1 minute to allow the agarose gel to dissolve. The mixture was then cooled to ca. 50-60 °C before adding 1 drop of ethidium bromide. The gel was cast in a tank with the required size comb and allowed to set under a fume hood for 1 hour. Once the gel set, the comb was removed and sufficient TBE buffer was added to ensure the gel was fully submerged. Samples were then loaded onto the gel, in the wells formed by the comb, and separated for 50 minutes by electrophoresis at 40 Volts. Imaging of the gel was under UV light, using the UVITech transilluminator (UVITech, Cambridge, UK).

2.3.3 DNA and RNA quality determination by gel electrophoresis

DNA and RNA were separated and analysed by electrophoresis on agarose gel stained with ethidium bromide to check their quality using the method described before in section 2.3.2. For both DNA and RNA was a 1.5% (w:v) agarose gel stained with ethidium bromide, 5 µL of nucleic acid samples were mixed with 2 µL of DNA loading dye (6x) (Thermo Fisher Scientific™, USA) and 5 µL of this mixture was loaded into the gel wells and separated. Both DNA and RNA were visualised under UV light as described in section 2.3.2.

2.3.4 Enzymatic manipulation of nucleic acids

2.3.4.1 Polymerase Chain Reaction (PCR)

PCR reactions were carried out in 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) using a final volume of 25 µL containing 0.2 µL (5 U/µL) of Taq DNA Polymerase and 2.5 µL of PCR buffer (F. Hoffmann-La Roche, Switzerland), gene specific forward and reverse primers (0.1 µL each primer) at a final concentration of 0.2 µM, 2.5 µL of dNTPs (Invitrogen™ Corporation, USA) at a final concentration of 0.2 mM and about 50 ng of DNA template (10 µL of DNA stock solution 1:10 dilution). The final volume of 25 µL was reached by adding 9.6 µL of sterile distilled water (SDW). Thermal cycling conditions were as follows: 5 minutes denaturation at 95 °C followed by 35 cycles of 94 °C for 1 minute, 57 °C (melting temperature for the primer pair) for 1 minute at 72 °C for the extension of the expected fragments.

PCR reagents	**volume (µL)**
dNTPs mix	250
reaction buffer	250
Taq polymerase	10
forward primer	20
reverse primer	20
sterile distilled wated (sdw)	960
[a]mix, in equal amount, of the four deoxynucleotides (i.e. dATP, dGTP, dCTP and dTTP).	

Table 2.1: Polymerase Chain Reaction (PCR) solution mix for 100 reactions.

Volume of the different reagents for the final of 1.5 mL suitable for 100 PCR reactions.

2.3.4.2 Enzymatic clean-up protocol for Sanger sequencing

5 µL of PCR reaction for each DNA sample were transferred to a new 0.2 mL non-skirted 96-well PCR plates. To each well, containing the amplified DNA, were added 1 µL of Shrimp Alkaline Phosphatase (rSAP) and 1 µL of Exonuclease I both from New England BioLabs Inc. (USA). The PCR plate was then moved to the thermocycler, for the second and last step of the clean-up protocol, with the following settings, 37 °C for 15 minutes and 80 °C for another 15 minutes; the latter temperature for the complete and irreversible inactivation of the enzymes.

2.3.4.3 Protocol for preparation of samples for fragment analysis (genotyping)

For 100 reactions, 10 µL of Gene Scan™ 500 ROX™ and 990 µL of highly deionized (Hi-Di) formamide both Applied Biosystems S.A. (USA) were mixed together. Two µL of SDW diluted PCR reaction were mixed with 8 µL of ROX mix reaction buffer and transferred to a new 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) for the fragment analysis (genotyping).

2.3.4.4 Design of primers

Gene specific primers for PCR were designed using the online software Primer3web version 4.1.0. (primer3.ut.ee). Primer sequences for Sanger sequencing were purchased from Eurofins Genomics (Ebersberg, Germany) while those primers for Fragment analysis (genotyping) were purchased from Sigma-Aldrich (USA) and of these, the forward primers were HEX (hexachloro-fluorescein) labelled.

2.3.4.5 Sanger sequencing of DNA

The fluorescent Sanger sequencing was carried out by the Genome Technology lab at The James Hutton Institute (Dundee, Scotland, UK). DNA was extracted (see section 2.2.2) from all the sixty-three different raspberry cultivars and selections (see Table A.1 appendix). Samples for the analysis were processed as described in sections 2.3.4.1 and 2.3.4.2; the only difference was the use of fluorescently labelled forward primers. Sequencing reactions were conducted on a capillary based Applied Biosystems AB3730 DNA Analyser (Applied Biosystems, USA) and raw data were collected using Data collection software version 4.0 and analysed using Applied Biosystem's Sequence Analysis version 6.0; all software Applied biosystem S.A. (USA).

2.3.4.6 Fragment analysis for genotyping

The DNA SSR genotyping was carried out by the Genome Technology lab at The James Hutton Institute (Dundee, Scotland, UK). DNA was extracted (see section 2.2.2) from all the sixty-three different raspberry cultivars and selections (see Table A.1 appendix). Samples for the analysis were processed as described in sections 2.3.4.1 and 2.3.4.3. The samples were processed with the Genetic Analyzer 3730 Applied Biosystems S.A. (USA), the data were collected by means of the Genetic Analyzer 3730 data collection software version 4.0 Applied Biosystems S.A. (USA). The final data analysis was performed with GeneMapper software version 5.0 Applied Biosystems™ (USA).

2.3.5 RNA microarray experiment set up

A *Rubus idaeus* microarray was previously developed by the Genome Technology lab at The James Hutton Institute (Dundee, Scotland, UK) using the Agilent platform (Agilent Technologies, USA). This custom microarray was designed from a unigene set assembled from existing sequence resources, comprising transcript sequences isolated from a range of *R. idaeus* tissues, developmental stages and conditions, including developing fruit and buds. This set was composed of sequences originating from four sources:

- Roche 454 transcripts (52,263)
- Illumina GAII transcripts (118,275)
- Sanger Expressed Sequence Tags (4,360)
- BAC coding sequences (1,425). In total, 176,833.

Sequences were assembled using CAP3 software (https://omictools.com/cap3-tool), generating 41,155 contigs and 22,098 singletons. These sequences were searched using BLASTx (https://omictools.com/blastx-tool) against known plant polypeptide sequences to identify the top protein homologues which, along with the presence of a polyA or polyT tract, enabled determination of the predicted orientation for 55,920 unigenes. Using eArray online software (Agilent Technologies Inc., USA) with default parameters, a total of 55,708 oligonucleotide probes (60mers) were designed and utilised for generation of a custom Agilent microarray in 8x 60k format (JHI_Ri_60k_v1).

RNA labelling as cDNA with Cy®3 fluorescent dye and downstream microarray processing were performed following manufacturer recommended protocol for the single-colour Microarray-Based Gene Expression Analysis, version 6.5 (Agilent Technologies, USA) and using the LOW Input Quick Amp Labelling Kit (Agilent Technologies, USA). A total of 24 microarrays were processed, consisting of four biological replicates for each of the three developmental stages (i.e. closed bud, open flower and green berry) per type (i.e. mostly and never 'crumbly'), see Table 2.2 for full details. Microarray scanning was performed with Agilent G2505B scanner (Agilent Technologies, USA), data were extracted from images using Agilent Feature Extraction software version 10.7.3.1 (Agilent Technologies, USA) and aligned with the appropriate array grid template file. The Feature Extraction (FE) datasets for each array were loaded into GeneSpring software, version 7.3, (Agilent Technologies, USA) for further analysis. Data were normalised using default double-channel settings: intensity values were set to a minimum of 0.01, data from each array were normalised to the 50th percentile on the array and the signal from each probe was subsequently normalised to the median of its value across all arrays. Unreliable data flagged as absent in 3/4 replicate samples were discarded the Feature Extraction software.

Sample	rep	stage	type	tube	slide	array
1	A	closed bud	mostly	1	1	1
2	A	closed bud	never	5		2
3	A	open flower	mostly	9		3
4	A	open flower	never	13		4
5	A	green berry	mostly	17		5
6	A	green berry	never	21		6
7	B	closed bud	mostly	2		7
8	B	closed bud	never	6		8
9	B	open flower	mostly	10	2	1
10	B	open flower	never	14		2
11	B	green berry	mostly	18		3
12	B	green berry	never	22		4
13	C	closed bud	mostly	3		5
14	C	closed bud	never	7		6
15	C	open flower	mostly	11		7
16	C	open flower	never	15		8
17	C	green berry	mostly	19	3	1
18	C	green berry	never	23		2
19	D	closed bud	mostly	4		3
20	D	closed bud	never	8		4
21	D	open flower	mostly	12		5
22	D	open flower	never	16		6
23	D	green berry	mostly	20		7
24	D	green berry	never	24		8

Table 2.2: 'Crumbly' microarray experiment layout.

Four slides, three development stages: close bud (**CB**), open flower (**OF**) and green berry (**GB**) for the two different plant types tested (i.e. mostly and never 'crumbly') and with four biological replicates (i.e. A, B, C and D).

2.4 Biochemical analysis

2.4.1 Phytohormones extraction

Plant material (both drupelets and receptacle from both green and red stages) was finely ground in liquid nitrogen with a mortar and pestle. The sample was then weighed (50 mg) into 2 mL Eppendorf tube and suspended in 1 mL of extraction solution, containing a mixture of isotopically labelled internal standards, and homogenised with 2 stainless steel beads (3 mm of diameter) onto a vibration mill, model MM 301 (Retsch GmbH & Co. KG, Haan, Germany) with an operating frequency of 30 Hz for 1 minute and homogenized. The mixture was first sonicated for 3 minutes in a cold room (4 °C) and then extracted using a benchtop laboratory rotator at 15 rpm for 30 min in cold room (4 °C). The extraction was followed by centrifugation at 16,600 g at 4 °C for 10 min, the supernatant was saved and transferred to a new 2 mL Eppendorf tube for further purification. Samples were purified using Oasis® HLB reverse-phase (C_{18}), polymer-based, solid phase extraction (RP-SPE) cartridge (1 cc/30 mg) from Waters Ltd (Elstree, UK). The solid phase extraction protocol consists of four steps. First, the cartridge was conditioned/activated with 1 mL of pure methanol followed by 1 mL of distilled water. The conditioning/activation was followed by the column equilibration step with 1 mL of 50% aqueous acetonitrile (v:v) and then by the sample loading. The flow through was collected in a 7 mL amber glass vial (SigmaAldrich Co. Ltd, UK) together with 1 mL of 30% aqueous acetonitrile solution (v:v) used to rinse the cartridge. Samples were dried in speed vacuum at 30 °C, for about 30 minutes to help remove all the organic solvent, soon after they were frozen in liquid nitrogen and then loaded on a freeze drier (Martin Christ Gefriertrocknungsanlagen GmbH, Germany) to be dried overnight. Once dried, the samples were briefly stored at -80 °C; before performing the analysis the samples were reconstituted with 40 µL of 30% aqueous acetonitrile (v:v) and analysed with the LC-MS/MS system.

2.4.2 QqQ-LC/MS

A Triple Quad LC/MS (Liquid Chromatography/Mass Spectrometry) Agilent 6460 Agilent Technologies™ was used for the targeted analysis of eighteen plant hormones (phytohormones). The Triple Quad LC/MS was equipped with the following components, Diode Array Detector (DAD), Thermostatted Column Compartment (TCC), Automatic Loading Sampler (ALS) and Quaternary Pump.

The sample extract (5 µL) was injected onto a 150 x 2 mm (5µm) Gemini RP C18 (110 Å) column fitted with a Gemini C18 4 x 2 mm Security guard cartridge (Phenomenex, Cheshire, UK). Samples were eluted at a flow rate of 0.3 ml min^{-1} using a gradient separation with two mobile phases A = 0.1% formic acid in deionized water and B = 0.1% formic acid in methanol. The elution gradient lasted 29 min in total and was as follows: A/B 95/5 (v/v) hold for 2 min, ramped up to 35% B in 1 min, followed by further ramping up from 35% to 55% in 15.5 min and further ramped up from 55% up to 100% in 3 min and hold for 2 min. Within 0.5 min the gradient was returned to the initial composition of 5% and held for 5 min.

All the individual twenty-seven individual metabolites, the eighteen target phytohormones and the nine isotopically labelled standards which were to be analysed on the QqQ-LC/MS system were optimised to generate fragment patterns which would allow for the generation of a multiple reaction monitoring (MRM) which is the process consisting of a series of steps: fragmentation, detection, quantification and monitoring of the target analytes peaks along the duration of the chromatography separation. Further details about the analytical method can be found on section 5.2 of chapter 5 of this work.

2.5 Statistics and bioinformatics

All statistical analyses were performed using Genstat version 18 (VSN International, UK) unless stated otherwise. All the details of the different analysis tests performed were reported in the experiment chapters (see chapter 3, 4 and 5).

2.5.1 Statistical analysis 'crumbly' microarray experiment

Statistically filtering of transcriptomic analysis was performed using ANOVA (Analysis Of VAriance, p-value <0.05) adjusted with Benjamin and Hochberg multiple testing correction. Transcript analysis was carried out in Genespring version 7.3 (Agilent Technologies, USA). The analysis of the microarray probes mapped inside the 'crumbly' QTLs (Quantitative Trait Loci) was performed in Genstat version 18 (VSN International, UK); see chapter 4 for further details.

2.5.2 Cluster Analysis

Cluster analysis was performed for several sets of transcriptome data with the aim of obtaining groups of probes linked based on their expression patterns. The 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* equivalent had ontology terms of interest or because the genes were mapped inside the 'crumbly' QTLs (see chapter 3 and 4 for details) were imported in Genespring version 7.3 (Agilent Technologies, USA) and the cluster analysis was performed using similarity measure of Pearson's correlation with clustering algorithm of average linkage.

2.5.3 Homology search

2.5.3.1 Blast search

Similarities between the *Rubus idaeus* genes from the Glen Moy genome assembly database (The James Hutton Institute - http://camel.hutton.ac.uk/raspberry/) and the *Arabidopsis thaliana* genes, were identified by using the Basic Local Alignment Search Tool (BLASTn) for nucleotide. The Glen Moy gene sequence, in FASTA format, from the Glen Moy genome assembly database were copied and pasted in The Arabidopsis Information Resource (TAIR - https://www.arabidopsis.org/). Matches were considered non-significant when e-value was greater than 0.01.

2.5.4 Gene Ontology (GO) annotations

In order to find insights into potential functions of the identified *Rubus idaeus* genes, the *Arabidopsis thaliana* genes, orthologs of those *R. idaeus* and matching microarray probes differently expressed, were analysed with the TAIR gene ontology bulk search tool (https://www.arabidopsis.org/tools/bulk/go/index.jsp). The *A. thaliana* gene IDs were loaded into the TAIR browser that provided the associated ontology terms for each gene.

2.6 Genomic analysis

2.6.1 Quantitative Trait Locus (QTL)

QTL positions were identified using a hidden Markov model (HMM) approach adapted from similar work for QTL mapping in autotetraploid species (Hackett et al., 2013), as initial QTL mapping using interval mapping in MapQTL 5 (Van Ooijen 2004) gave logarithm of the odds (LOD) profiles that were unexpectedly irregular, given the high-density map, resulting in uncertainty in locating the peak LOD score. The LOD score is the usual LOD score for QTL mapping (i.e. $\log_{10}$ likelihood of QTL at that position/likelihood of no QTL) and its significance was tested using a permutation test based on 200 permutations that allowed to estimate the genome-wide threshold, for a significance level of 0.05, equivalent to LOD = 3.7 (Hackett et al., 2018). The point estimation of the QTL location, the position along the map with the highest LOD for that specific QTL, was estimated by means of the 'LOD drop-off method' that corresponds to a decrease of one from the maximum LOD score (Hackett, 2002).

2.6.2 SNPs detection

The amplified fragments regions of the five 'crumbly' markers mapped inside the three 'crumbly' QTLs were sequenced with the procedure described before (see section 2.3.4.2). Three of the markers were target genes, located in coding regions while the other 2 were gene 'tags', located in non-coding regions of the genome. The sequences were imported into Sequencer software version 5.4.6 (Gene Code Corporation, USA), assembled using the automatic function and the produced contig was manually inspected to identify any possible SNP (Single Nucleotide Polymorphism) along the assembled sequence.

2.7 Linkage groups drawing

Linkage groups and associated QTLs were produced by importing the list of markers and their corresponding position in (cM) into MapChart software version 2.32 (Wageningen University, The Netherlands).

Chapter 3: Gene expression analysis between mostly ‘crumbly’ vs. never ‘crumbly’ phenotypes from the Glen Moy x Latham progeny

3.1 Introduction

Despite botanical differences, all fruits undergo similar developmental steps of fruit set, growth, maturation and ripening. Fruit set is the first stage of development after fertilisation. It is followed by an active cell division and expansion phase, called growth during which the fruit attains its maximum size. This in turn is followed by ripening, the last of its development stages (Kumar et al., 2014). The current model of regular fruit set implies that ovary growth is blocked before pollination and that auxin is a key regulator of ovary growth de-repression at fruit set. Other phytohormones have been shown to have a role in fruit initiation and development (i.e. gibberellin, cytokinin, brassinosteroids, ethylene and abscisic acid) (Goetz et al., 2007, Pandolfini et al., 2007).

Plant hormones with their unquestionable role (Ozga and Reinecke, 2003, Pandolfini et al., 2007, Nagpal et al., 2005, Kumar et al., 2014, Vanstraelen and Benkova, 2012) in regulating all the processes related to fruit development, as well as the molecular processes behind the development of the flower through fertilization of the ovules to fruit development are all important aspects of study in raspberry because they could potentially bring to light the mechanisms responsible for the formation of 'crumbly' like misshapen fruit in red raspberry.

Transcriptomics technologies are techniques that are used to study an organism transcriptome (sum of all RNA transcripts) allowing gene expression analysis to identify genes whose expression levels vary due to specific treatments, different developmental stages, between different tissues etc. RNA-microarray technology is a tool containing thousands of DNA fragments (probes) of known sequence arrayed on a chip. This easily allows the collection of thousands of data points related to levels of gene expression between samples in formats that can be used for bioinformatic and statistical analysis (Waters et al., 2005). Microarray analysis has been applied with success to gene expression analyses of different plant species (i.e. tomato, strawberry, peach, pear and grape) during fruit development, in particular fruit ripening, allowing the identification of genes specifically involved in different stages of fruit growth (Waters et al., 2005). It constitutes a relatively easy and convenient way to select, for instance, potential genes of interest for breeding programmes (Slater et al., 2008).

In this chapter the use of microarray analysis is described to identify variations in gene expression in samples of red raspberry from three different developmental stages (i.e.

closed bud, open flower and green berry) and for two different phenotypes (i.e. mostly and never ‘crumbly’).

3.2 Materials and methods

RNA extraction is complicated from samples like raspberry containing high levels of polysaccharides, polyphenols and protein contaminants (Jones et al., 1997). Total RNA was extracted using the protocol described in material & methods (see section 2.2.1). RNA quality and its concentration were estimated spectrophotometrically from all samples using a NanoDrop™ spectrophotometer (NanoDrop, USA). This tested the RNA for potential contaminants (i.e. proteins, phenols, carbohydrates etc.) and provided an estimate of the RNA concentration in ng/µL. Two ratios were determined, 260/280 and 230/260 nm. A ratio 260/280 of about 2.0 indicated a 'pure' RNA sample while lower ratios indicated that samples were potentially contaminated with proteins and/or phenols absorbing at 280 nm. The second ratio 260/230 gave a further quality check of the RNA purity with a 2.0-2.2 ratio indicating good quality RNA, while lower ratios were signals for contamination by compounds absorbing at 230 nm such as carbohydrates. The ratios range values of the raspberry samples were between 2.18 and 1.95 for the 260/280 ratio and between 2.13 and 1.57 for the 260/230 ratio (see Table 3.1 for more details), indicating good quality of RNA.

Once the RNA quality had been assessed, 6 µL of each sample were run through an electrophoresis gel (see section 2.3.2 for further details). The two clear bands of ribosomal RNA (rRNA) 18S and 28S indicated good quality samples. Before proceeding with the microarray analysis, the quality of the RNA samples was further examined using the RNA Integrity Number (RIN) algorithm with a 2100 Bioanalyzer system (Agilent, USA) that analysed, after electrophoretic separation, the ribosomal RNA of the samples. Essentially this instrument gave a more reliable and accurate measure of RNA integrity; RIN values of ten stated intact RNA while RIN values of 1 indicated completely degraded RNA (Schroder et al., 2006). The RIN range values of the raspberry samples were between 8.90 to 9.90 indicating a very good quality RNA material, see Table 3.1 for further details.

Sample	sample ID	260/280 nm	260/230 nm	ng/µL	RIN
closed bud_mostly	1	2.15	1.97	1550.17	9.70
	2	2.15	2.06	1651.69	9.70
	3	2.15	1.75	1187.78	9.50
	4	2.07	1.85	1245.59	9.70
closed bud_never	5	2.15	1.76	1478.47	9.60
	6	2.14	2.13	1811	9.60
	7	2.16	1.78	1484.50	9.60
	8	2.18	1.97	1370.97	9.70
open flower_mostly	9	2.03	1.67	681.30	9.50
	10	2.03	1.72	726.87	9.40
	11	1.99	1.59	746.54	9.40
	12	1.99	1.60	959.29	9.50
open flower_never	13	1.99	1.72	638.79	8.90
	14	1.95	1.57	904.03	9.80
	15	1.96	1.58	747	9.90
	16	2.01	1.72	894.09	9.90
green berry_mostly	17	2.03	1.89	445.87	9.40
	18	1.99	1.58	492.87	9.40
	19	2.04	1.68	441.17	9.40
	20	2.00	1.61	472.65	9.60
green berry_never	21	2.07	1.93	389.11	9.00
	22	1.99	1.65	640.43	9.50
	23	1.99	1.67	512.83	9.60
	24	2.00	1.61	535.70	9.50
target		1.5-2	>1.5		0-10

Table 3.1: RNA quality indicators for the 24 samples of the 'crumbly' microarray.

Quality indicators of RNA samples extracted from three different stages (i.e. (**CB**) closed bud, (**OF**) open flower and (**GB**) green berry) of two different treatments (i.e. mostly and never 'crumbly') and target values for spectrometric NanoDrop™ analysis, ratios for 260/280 nm and 260/230 nm, yield in ng/µL and RIN (Bioanalyzer) number at a scale 0-10.

3.2.1 Experimental design and data analysis

The *Rubus idaeus* microarray was previously developed by the Genome Technology lab at The James Hutton Institute (Dundee, Scotland, UK) using the Agilent platform (Agilent Technologies, USA) see section 2.3.5 of materials and methods chapter for further details. The microarray contains in total 55,708 single 60-mer oligonucleotide probes representing unique transcripts. The samples analysed consisted of 14 individual progeny from a Glen Moy x Latham mapping population of which seven were labelled as mostly 'crumbly' phenotype since they consistently produced, across many years of scoring, primarily fruits with uneven shape (though not in every season thus the mostly crumbly designation), while the other seven plants were labelled as never 'crumbly' phenotype since they always produced regular shape fruits.

For each plant and phenotype, three different development stages were examined, closed bud (CB), open flower (OF) and green berry (GB). Four biological replicates were collected for each stage giving a total of 168 samples from which the RNA was extracted and processed. For the microarray analysis 24 pools were created by merging the seven samples for each of the sample typology (e.g. CB_mostly, CB_never, OF_mostly, etc.), repeating the procedure for each of the four biological replicate available for each plant. In Table 3.2, as example, were reported all the samples needed to produce the 4 pools related to the CB mostly 'crumbly' sample. This specific design was chosen in order to reduce the effect of the environment and genetic differences not associated with the 'crumbly' trait on the expression level.

In total 24 microarrays were processed, data were extracted using Feature Extraction software (Agilent Technologies, USA) and then imported into GeneSpring software (version 7.3) (Agilent Technologies, USA) for data pre-processing and normalization. To make sure labelling differences were considered, the 1-colour normalization with the default setting was performed. All the probes with signal indistinguishable or too small to be significantly different from the background signal were removed from the data. After filtering, data were normalised (i.e. transformed in logarithm on base 2), the remaining probes were subjected to statistical analysis (i.e. two-way ANOVA) with stage (i.e. closed bud, open flower and green berry) and phenotype (i.e. mostly and never 'crumbly') as factors, the analysis was performed using GeneSpring software (version 7.3) (Agilent Technologies, USA) to identify genes with significant differential expression dependent on the two factors and their interactions. A stringency threshold, p-

value ≤0.05 was applied for the analysis and from these, 827 probes exhibiting significant changes in the expression levels were identified. The same two-way ANOVA was repeated in GenStat (VSN International, UK) too and this because this statistical package provide in the results of the analysis all the figures (i.e. predicted means, standard error, etc.) that were used to produce Tables to furnish even with numerical figures the results of the analysis and give to the reader the complete overview of the situation.

Replicate	Rep A	Rep B	Rep C	Rep D
sample	1A_CB_mostly	1B_CB_mostly	1C_CB_mostly	1D_CB_mostly
	2A_CB_mostly	2B_CB_mostly	2C_CB_mostly	2D_CB_mostly
	3A_CB_mostly	3B_CB_mostly	3C_CB_mostly	3D_CB_mostly
	4A_CB_mostly	4B_CB_mostly	4C_CB_mostly	4D_CB_mostly
	5A_CB_mostly	5B_CB_mostly	5C_CB_mostly	5D_CB_mostly
	6A_CB_mostly	6B_CB_mostly	6C_CB_mostly	6D_CB_mostly
	7A_CB_mostly	7B_CB_mostly	7C_CB_mostly	7D_CB_mostly
pool	pool A (1A+2A…+7A)	pool B (1B+2B…+7B)	pool C (1C+2C…+7C)	pool D (1D+2D…+7D)

Table 3.2: Samples pooling relative to the closed bud (CB) of the mostly 'crumbly' plants.
The four biological replicates collected from each of the seven plants were named: A, B, C and D. All the seven replicates named A were pooled together as showed in the last row of column (**Rep A**), the same applied to the seven replicates B, C and D. This produced the four pool A, B, C and D that formed the 4 reps for the mostly 'crumbly' closed bud samples of the 'crumbly' microarray experiment. The same procedure was repeated for the other two stages, open flower (OF) and green berry (GB).

3.2.2 Heatmap tree clustering

The analysis was performed using GeneSpring software (v. 7.3) (Agilent Technologies, USA) using similarity measure of Pearson's correlation with clustering algorithm of average linkage. The analysis was then focused only on a limited number of probes, 107 in total (see Table A.3.1 in appendix for full list of probes), those matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had specific gene ontology annotations related to: flower development, hormones, pollen and transport.

3.2.3 Gene Ontology (GO) term annotation

The probes differently expressed between the two different phenotypes (i.e. mostly and never 'crumbly') were first analysed through the Glen Moy genomic assembly browser (http://camel.hutton.ac.uk/raspberry/) to identify the correct matching *Rubus idaeus* gene. For each identified *R. idaeus* gene, the browser reported the corresponding *Arabidopsis thaliana* ortholog gene and to make sure the correspondence of the genes between the two different species was correct, the full length genomic sequence (https://www.arabidopsis.org) was copied and then blasted in the Glen Moy genomic assembly browser to ensure that the selected *A. thaliana* was the correct ortholog of the *R. idaeus* corresponding gene. The selected genes were then analysed through (https://www.arabidopsis.org/tools/bulk/go), the *Arabidopsis thaliana* online browser that display in detail all the Gene Ontology (GO) annotations done to each selected gene.

3.3 Results

Heatmaps represent one of the easiest ways to obtain a clear overview of the results of large data set such as those of microarray analysis. They allow graphical representation of complex data in which values are depicted by colours. Heatmaps make it easy to visualize complex data that are organized and displayed according to similarity in their expression levels. Such approach creates data patterns in matrices that allow the visualisation of differences in the expression levels between the different samples. Upregulated genes are depicted as red marks while downregulated genes as blue marks in a red to blue-scale shading, creating a matrix with colour patterns that allow the distinguishing of clusters of probes matching genes with clear differences in the expression levels between the samples of an experiment. Here the heatmaps were used to identify clusters of probes differently expressed between two phenotypes (i.e. mostly and never 'crumbly') for each of the three development stages studied (i.e. closed bud, open flower and green berry).

The 827 probes, selected after statistical analysis (see section 2.5.1), differently expressed between the two phenotypes, with difference being significant at 95% confidence levels were subjected to a heatmap tree clustering analysis (section 2.5.2). The resulting heatmap was reported in Figure 3.1 and two main clusters were highlighted, (curly brackets) and named **A** and **B**. The first cluster showed a specific pattern of probes being upregulated in the mostly 'crumbly' and downregulated in the never 'crumbly' phenotype for all the three different development stages (i.e. closed bud, open flower and green berry). On the other hand, cluster **B** showed a situation opposite to cluster **A**, here probes were downregulated in the mostly 'crumbly' compared to the never where probes were downregulated. Together with these two big clusters (i.e. **A** and **B**), a further small cluster of 25 probes, this time showing differences in the expression level only at open flower stage, was identified and highlighted with a black circle in the heatmap (see Figure 3.1).

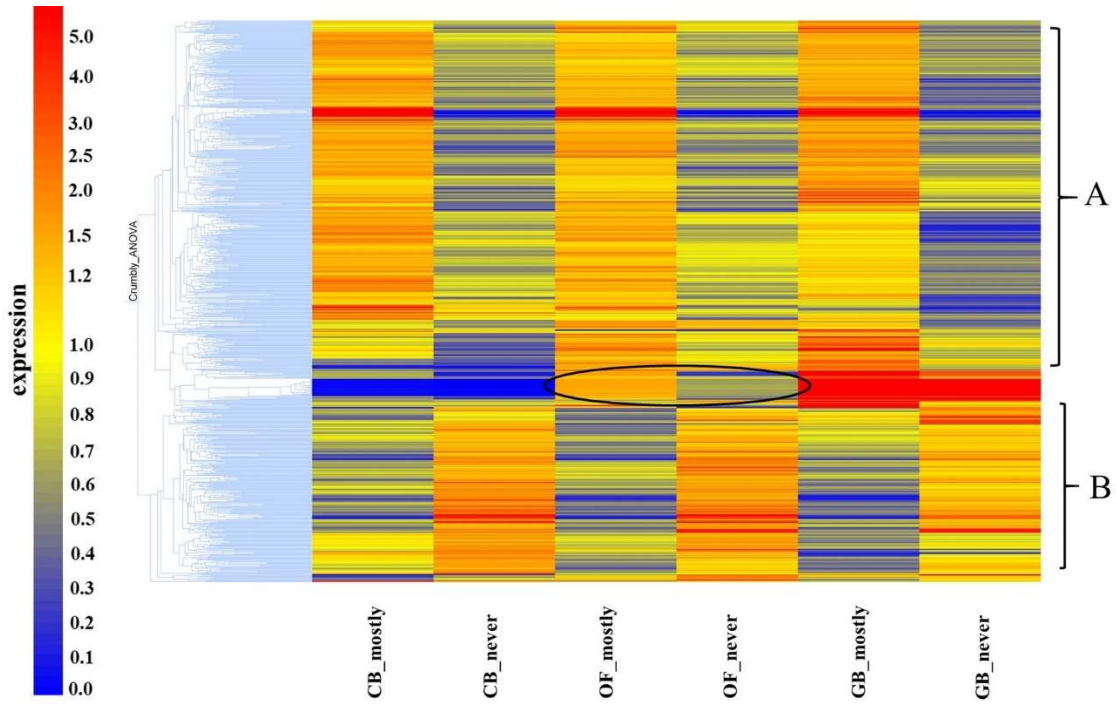


Figure 3.1: Tree cluster heatmap of the microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly').

Gene-tree cluster heatmap of the 827 genes differently expressed between the two different phenotypes (i.e. mostly and never 'crumbly') for all three development stages (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**); the difference in the expression level of the microarray probes being significative at 95% confidence levels. Two main big clusters of genes were highlighted by the heatmaps; they were named **A** and **B** respectively with the first showing genes upregulation in the mostly 'crumbly' phenotype samples and downregulated in the never 'crumbly' phenotype samples while the **B** cluster showed genes downregulated in the mostly 'crumbly' and upregulated in the never 'crumbly' phenotype in all the three different stages. A further small cluster (highlighted in the picture with a black circle), showed 25 probes matching as many genes whose expression levels changed drastically between the two phenotypes but only at the open flower stage. High expression levels are indicated in red colour while low expression levels in blue colour as per scale bar presented on the left side of the heatmap.

The three clusters highlighted in Figure 3.1 contained about 500 probes and to select only those matching genes with potential functions related to 'crumbly' fruit, a further screening was performed. All the *Arabidopsis thaliana* ortholog genes of those matched by the selected probes were analysed one by one with (https://www.arabidopsis.org), the *Arabidopsis thaliana* browser for gene ontology annotations. The analysis was focused on four different annotation terms: pollen, flower development, hormones and transport (http://geneontology.org/). The choice to target primarily these four specific factors was driven by suggestion in the literature that in some varieties (e.g. Sumner) the 'crumbly' fruit phenotype might be caused by mutations producing the homozygous state for two

recessive pairs of genes that retard the development of the embryo sac and reduce the production of fertile pollen (Jennings, 1988). This suggested that those genes having functions that affect pollen production and functions such as those involved in flower development must be considered since they might be directly involved in processes and/or functions potentially causing 'crumbly' fruit. Hormones are the main plant growth regulators and thus their involvement cannot be excluded. Gene ontology terms related to transport were chosen also because, according to the hypothesis behind this work, a crosstalk between receptacle and fertilized ovaries would be needed to guarantee the synchronised growth of all the ovaries that form the final fruit. Without such a regulating process, mediated by the receptacle, the late fertilised ovary would never be able to reach the same size of those fertilised earlier with potential consequence of misshapen fruits.

The gene ontology analysis on the microarray probes from the three selected clusters highlighted in heatmap of Figure 3.1, allowed the selection of 107 probes, all differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and all having gene ontology annotations related to pollen and/or flower development, and/or hormones and/or transport.

A new tree cluster heatmap was produced, using the same procedure described in section 2.5.2 but this time only the 107 microarray probes matching genes with gene ontology annotations related to flower development, hormones, pollen and transport were imported in gene spring. The tree cluster heatmap for the 107 probes (Figure 3.2) displayed two main clusters with respectively 37 (cluster **C**) and 36 (cluster **D**) probes. Interestingly all the four different classes of gene ontology terms (i.e. flower development, hormones, pollen and transport) were represented in these two cluster. Further screening of the 107 probes was performed in order to select probes and their corresponding genes specific to each of the gene ontology terms (i.e. flower development, hormones, pollen and transport).

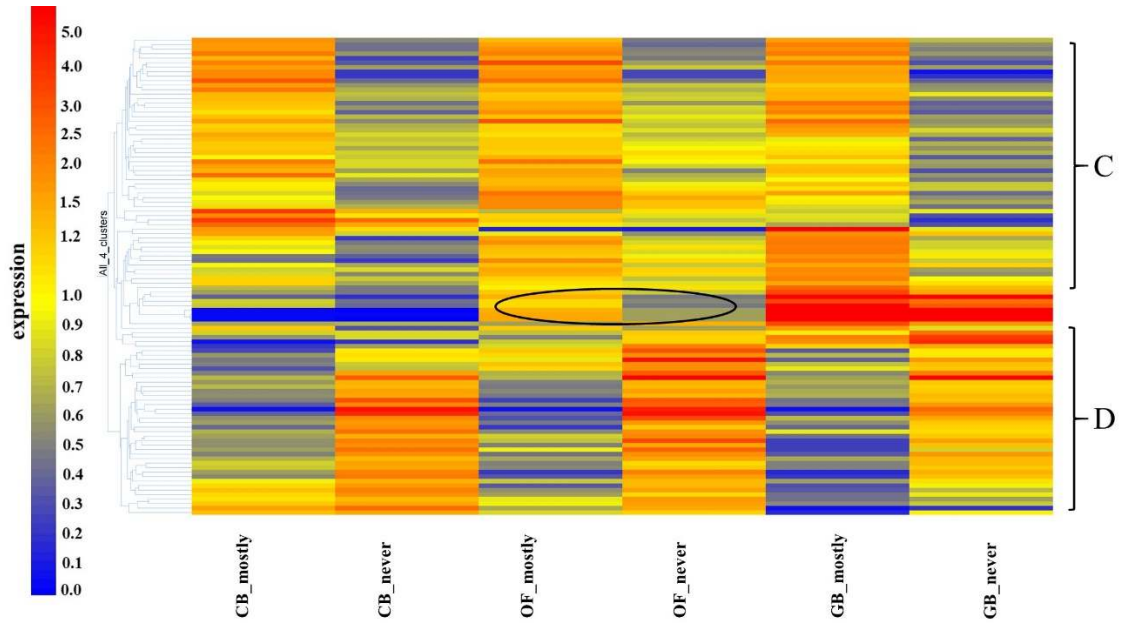


Figure 3.2: Tree cluster heatmap of the microarray probes differently expressed and matching genes with gene ontology annotation related to: flower development, hormones, pollen and transport.
Tree clustering heatmap of the 107 microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and for the three different development stages (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**). Two main clusters of differently expressed probes were highlighted. They were named **C** and **D** and they contained respectively 38 and 37 probes (75 in total) which matched as many Arabidopsis thaliana ortholog genes. The probes of cluster **C** were upregulated in the mostly 'crumbly' compared to never one while the opposite for the probes within the cluster **D**. High expression levels are indicated in red colour while low expression levels in blue colour as per scale bar presented on the left side of the heatmap.

3.3.1 Heatmap of microarray probes matching genes with GO terms related to pollen

Twenty-one probes matched *Rubus idaeus* genes whose *Arabidopis thaliana* counterparts had gene ontology annotations related to pollen. These twenty-one probes were imported into GeneSpring software and following the procedure described in section 2.5.2, a tree cluster heatmap, specific for these 21 probes, was created (Figure 3.3). The full list of the 13 probes belonging to the two clusters highlighted in the tree cluster heatmap of Figure 3.3 was reported on Table 3.3 while the list for the remaining 8 probes, those outside the two clusters of Figure 3.3, was reported in Table A.3.2 in appendix .

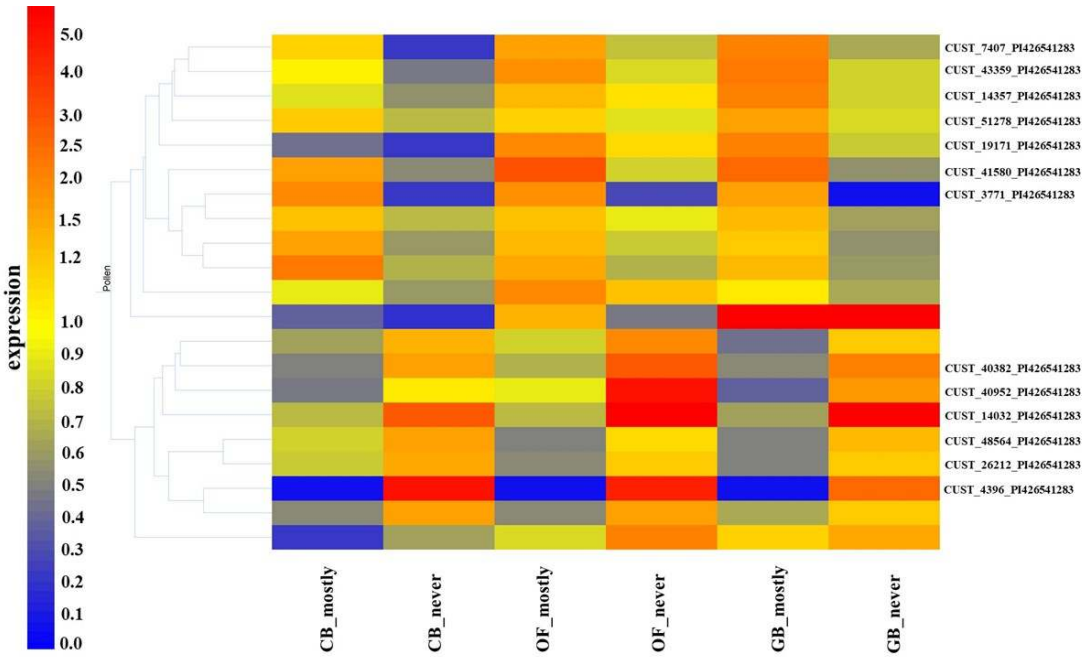


Figure 3.3: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to pollen.

Cluster tree heatmaps for the twenty-one microarray probes differently expressed for what concerned the stage*phenotype interaction with differences being significative at 95% confidence levels. Three stages tested (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**)) and two phenotypes (i.e. mostly and never 'crumbly'). The twenty-one probes matched *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to pollen. Two main clusters of probes with evident difference in the expression level, as highlighted by the colour shades of the heatmap, were selected and the corresponding probes were reported on the right side of the heatmap. With respect to the first two stages (i.e. closed bud and open flower), the seven probes within the first highlighted cluster were upregulated in the mostly 'crumbly' phenotype compared to the never one while the six probes within the second cluster were downregulated in the mostly 'crumbly' phenotype and upregulated in the never 'crumbly'. High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented.

The data of interest, in respect to genes related to pollen, were primarily those related to open flower stage and secondarily to those at closed bud stage because by green berry stage the drupelets were already set, thus the study of the expression levels for the genes controlling any function related to pollen was less relevant. In Tables 3.4 and 3.5, the *Arabidopsis thaliana* gene IDs for the 12 *Arabidopsis thaliana* genes, whose ortholog *Rubus idaeus*' one, were matched by the 'crumbly' microarray experiment probes were reported together with the predicted means with their standard error from the analysis of variance (ANOVA) with a stage (i.e. closed bud, open flower and green berry) per phenotype (i.e. mostly and never 'crumbly') interaction.

	Microarray probe	Gene Ontology	GO code	*A. thaliana* gene ID
1st cluster	CUST_7407_PI426541283	rejection of self-pollen	GO:0060320	AT4G24973.1
	CUST_43359_PI426541283	pollen tube development	GO:0048868	AT3G01640.1
	CUST_14357_PI426541283	recognition of pollen	GO:0048544	AT1G32300.1
	CUST_51278_PI426541283	pollen tube guidance	GO:0010183	AT2G15890.1
	CUST_19171_PI426541283	pollen development	GO:0009555	AT1G63180.1
	CUST_41580_PI426541283	pollen tube growth	GO:0009860	AT2G36880.2
	CUST_3771_PI426541283	regulation of pollen tube growth	GO:0080092	AT2G13680.1
		pollen germination	GO:0009846	
		microsporogenesis	GO:0009556	
		pollen wall assembly	GO:0010208	
2nd cluster	CUST_40382_PI426541283	pollen tube reception	GO:0010483	AT3G51550.1
	CUST_40952_PI426541283	pollen development	GO:0009555	AT5G15650.1
	CUST_14032_PI426541283	anther wall tapetum development	GO:0048658	AT2G31220.1
		pollen development	GO:0009555	
	CUST_48564_PI426541283	pollen tube growth	GO:0009860	AT1G63530.1
	CUST_26212_PI426541283	anther development	GO:0048653	AT4G30520.1
	CUST_4396_PI426541283	pollen development	GO:0009555	AT1G42470.1

Table 3.3: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to pollen and belonging to the two clusters of probes in tree cluster heatmap of Figure 3.3.

List of the thirteen microarray probes, significantly differently expressed for what concerned the stage*phenotype interaction with differences being significant at 95% confidence levels. The gene ID of the A. thaliana gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to pollen. In the first column of the table, the two clusters of the heatmap to which the probes belong were indicated.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0060320 (rejection of self-pollen)	AT4G24973.1	mostly	0.063	0.214	0.357	0.0848	CUST_7407_PI426541283
		never	-0.653	-0.125	-0.175		
GO:0048868 (pollen tube development)	AT3G01640.1	mostly	0.005	0.296	0.362	0.0538	CUST_43359_PI426541283
		never	-0.312	-0.061	-0.087		
GO:0048544 (recognition of pollen)	AT1G32300.1	mostly	-0.339	0.318	0.33	0.0725	CUST_14357_PI426541283
		never	-0.624	0.04	-0.092		
GO:0010183 (pollen tube guidance)	AT2G15890.1	mostly	-0.058	0.127	0.358	0.0487	CUST_51278_PI426541283
		never	-0.232	0.019	-0.091		
GO:0009555 (pollen development)	AT1G63180.1	mostly	0.08	0.061	0.221	0.0503	CUST_19171_PI426541283
		never	-0.142	-0.05	-0.073		
GO:0009860 (pollen tube growth)	AT2G36880.2	mostly	-0.427	0.17	1.496	0.081	CUST_41580_PI426541283
		never	-0.744	-0.316	0.985		
[b]GO:0080092 (reg. of poll. tube gr.) GO:0009846 (pollen germination) GO:0009556 (microsporogenesis) GO:0010208 (pollen wall assembly)	AT2G13680.1	mostly	0.224	0.515	0.43	0.1144	CUST_3771_PI426541283
		never	-0.26	-0.087	-0.25		

[a]2 degrees of freedom and 4 replicates
[b]GO:0080092 (regulation of pollen tube growth)

Table 3.4: ANOVA table of means for all the probes within the bottom cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to pollen.

ANOVA means of the normalised (i.e. logarithm of base 2) expression levels for the interaction, phenotype per stage, for the six probes within the second cluster highlighted in the pollen genes heatmap (Figure 3.3). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences, if any, in the expression levels of the probes were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry) were reported. For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations related to pollen were reported.

Gene Ontology (GO) term	A. thaliana gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0010483 (pollen tube reception)	AT3G51550.1	mostly	-0.295	-0.158	-0.262	0.1306	CUST_40382_PI426541283
		never	0.21	0.481	0.355		
GO:0009555 (pollen development)	AT5G15650.1	mostly	-0.318	-0.03	-0.431	0.1046	CUST_40952_PI426541283
		never	0.018	0.728	0.251		
[b]GO:0048658 (anther wall tap. dev.) GO:0009555 (pollen development)	AT2G31220.1	mostly	-0.146	-0.131	-0.201	0.1114	CUST_14032_PI426541283
		never	0.489	1.596	1.899		
GO:0009860 (pollen tube growth)	AT1G63530.1	mostly	-0.087	-0.29	-0.288	0.0432	CUST_48564_PI426541283
		never	0.212	0.042	0.144		
GO:0048653 (anther development)	AT4G30520.1	mostly	-0.097	-0.279	-0.296	0.0579	CUST_26212_PI426541283
		never	0.192	0.075	0.088		
GO:0048653 (pollen development)	AT1G42470.1	mostly	-1.263	-1.226	-1.302	0.0647	CUST_4396_PI426541283
		never	0.734	0.683	0.445		

[a]2 degrees of freedom and 4 replicates
[b]GO:0048658 (anther wall tapetum development)

Table 3.5: ANOVA table of means for all the probes within the bottom cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to pollen.

ANOVA means of the normalised (i.e. logarithm of base 2) expression levels for the interaction, phenotype per stage, for the six probes within the second cluster highlighted in the pollen genes heatmap (Figure 3.3). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences, if any, in the expression levels of the probes were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry) were reported. For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations related to pollen were reported.

For the seven *Arabidopsis thaliana* genes, matched by probes within the first cluster of the heatmap (Figure 3.3), nine different gene ontology terms related to pollen were associated with them and going into details, the majority of them (Table 3.2) indicated a potential effect of these genes in the processes linked with recognition of pollen, its germination and subsequent growth and development of the pollen tube inside the style and ovary of the flower. It would seem that the highlighted cluster of seven genes was related to the initial process of flower fertilization, from when the pollen reach the stigma of the carpels, passing to the step of its recognition/acceptance by the flower, its germination and until the growth and development of the pollen tube inside the style of the carpel.

The function of the seven genes identified in the first cluster of the heatmap in Figure 3.3 were:

- AT4G24973.1 – encodes a plant self-incompatibility protein S1 family with unknown function (source TAIR https://ui.arabidopsis.org/) and its higher expression level might cause the rejection of pollen by the stigmas of the same flower and/or plant with consequent reduced number of fertilised ovaries and potential development of misshapen fruit with lower number of drupelets.
- AT3G01640.1 – encodes a GLUCURONOKINASE G that phosphorylates D-GlcA to D-GlcA-1-phosphate. The enzyme is involved in the synthesis of UDP-glucuronate, (UDP: uridine diphosphate) a nucleotide sugar acting as glycosyl donor in the process of biosynthesis of carbohydrate polymers and glycoproteins that together compose the cell wall. Glucuronokinase is a novel member of the GHMP-kinase superfamily having a unique substrate specificity for d-glucuronic acid. In *Arabidopsis thaliana* the gene is expressed in all plant tissues with a preference for pollen where it contributes to providing the cell wall polymers needed during the development of the pollen tube (Pieslinger et al., 2010).
- AT1G34300.1 – encodes an enzyme belonging to the lectin protein kinase family involved in the process of pollen recognition. This gene was upregulated in the mostly 'crumbly' phenotype (source TAIR https://ui.arabidopsis.org/).
- AT2G15890.1 – encodes CBP1 that is involved in the activation of transcription factors responsible for the activation of the molecular processes that regulate the pollen tube attraction in the central cell of the ovule (Li et al. 2015).
- AT1G63180.1 – encodes an enzyme with UDP-D-glucose 4-epimerase activity. The protein catalyses the reversible conversion of UDP-galactose to UDP-glucose

(Seifert, 2004). It is involved in pollen development (source TAIR https://ui.arabidopsis.org/). The nucleotide sugar UDP-glucuronate, is a glycosyl donor in the process of biosynthesis of cell wall components such as carbohydrate polymers and glycoproteins (Pieslinger et al., 2010).

- AT2G36880.2 - encodes for MAT3 one of the four SAMs of *Arabidopsis thaliana*. S-Adenosylmethionine synthetase (SAMS), also known as methionine adenosyltransferase (MAT), an enzyme that synthesizes S-adenosylmethionine (SAM) from ATP and l-Met (Binet et al., 2011). SAM is the main enzyme involved in methionine metabolism as well as participating in the processes for the production of the precursors of ethylene and polyamine; with these last being important for pollen germination and pollen tube growth. MAT3 is expressed predominantly in pollen and it is required for pollen germination and pollen tube growth (Chen et al., 2016)
- AT2G13680.1 – encodes for a CALLOSE SYNTHASE 5 (CalS5 or GLS2) an enzyme with glucosyltransferase activity responsible for the synthesis of callose. CalS5 belongs to the first group of callose synthetase, responsible for fertility and cell division. CalS5 is specifically expressed in anthers, in microspores and pollen. The enzyme plays an important role in microgametogenesis and in fact it might be evolved as a key enzyme in the callose synthesis for both pollen development and pollen-tube growth. Three main functions have been identified for CalS5, patterning the exine layer, forming callose in pollen tubes, and preventing pollen degeneration at the early stages of pollen development (Zaveska Drabkova and Honys, 2017).

An overview on the molecular functions conducted by the seven genes included in the first cluster highlighted on Figure 3.3 showed that three genes (i.e. AT3G01640.1, AT1G63180.1 and AT2G13680.1) appear to be involved and contributing to pollen formation and development, in particular they all play a role in the formation of the cell wall of the pollen grains. Two genes (AT4G24973.1 and AT1G34300.1) are involved respectively in the rejection of the self-pollen and in general in its recognition by cells in the stigma of the carpels; such genes are then involved in the early steps of the fertilization, soon after the pollen reaches the stigmas of the carpels. The last to genes, AT2G36880.2 and AT2G15890.1, are involved respectively in part of the processes controlling pollen germination and the growth of the pollen tube as well as part of those regulating the attraction of the pollen tube to the central cell in the ovule. By focusing on the open flower stage, the sole relevant in relation to the biological/physiological

processes where pollen is involved. The seven *Arabidopsis thaliana* genes, orthologs of the seven *Rubus idaeus* genes matched by the microarray primers, were all upregulated in the mostly 'crumbly' phenotype compared to the never 'crumbly' one. This would suggest that overexpression of these seven genes might disrupt normal functioning and contribute to the formation of 'crumbly' like misshapen fruit in red raspberry.

The second cluster of probes, highlighted at the bottom of the tree cluster heatmap of Figure 3.3, contained six genes whose gene ontology annotations, in respect to pollen, referenced: pollen development (three genes), development of the anther (two genes) the process of flower fertilization and more precisely to pollen tube acceptance (one gene). This second cluster would represent the processes related to the formation of pollen, and here the closed bud stage is probably the most important since both anther and pollen development should begin at this development stage.

The function of the six genes identified in the second cluster of the heatmap in Figure 3.3 were:

- AT3G51550.1 – encodes for a receptor like kinase (RLK) that may play a crucial role in the pollen tube and ovule interaction, crucial for the fertilization process (Haruta et al., 2018a).
- AT5G15650.1 - RGP2 is a UDP-arabinose mutase that catalyses the interconversion between the pyranose and furanose forms of UDP-L-arabinose. It appears to be required for proper cell wall formation. RGPs are localised in Golgi apparatus and are auto-glycosylate with various uridine diphosphate (UDP)-sugars. Because of these two features, RGPs are thought to play roles in polysaccharide metabolism, particularly in cell wall and starch synthesis. RGP2 is highly expressed in actively dividing tissues such as pollen microspore (Burch-Smith et al., 2012).
- AT2G31220.1 – encodes for the Basic Helix Loop Helix protein 10 (bHLH010), in *Arabidopsis thaliana*, a transcription factor that together with bHLH089 and bHLH091 is important for the normal transcriptome of the developing anther (source TAIR: https://ui.arabidopsis.org/).
- AT1G63530.1 – encodes a hypothetical protein located in the nucleus with unknown function involved in pollen tube development (source TAIR: https://ui.arabidopsis.org/). Pollen tubes germinating on and elongating through the carpel style grow faster than those germinating on an artificial medium. Comparison of the expression levels, of pollen grains germinating on flowers with those on artificial medium, showed distinctive expression profiles with the pollen germinating

in vivo expressing a larger part of its genome compared to the pollen grown on artificial medium. Such findings suggest, for the interaction pollen tube/carpel, the activation of a regulatory network that orchestrates gene expression as pollen tubes migrate through the carpel (Qin et al., 2009).

- AT4G30520.1 – encodes SARK (SENESCENCE-ASSOCIATED RECEPTOR-LIKE KINASE). This regulates leaf senescence through synergistic actions of auxin and ethylene. SARK is involved in the development of the anther and is expressed in carpels and in general in flowers (source TAIR: https://ui.arabidopsis.org/).
- AT1G42470.1 - encoding an ortholog putative Niemann-Pick C1 protein (AtNPC1) which belongs to a class of protein containing a sterol-sensing domain (SSD) could be involved in the regulation of sterol pathway (Feldman et al., 2015).

The results of the analysis of the function, of the six genes within this second cluster of Figure 3.3, showed that AT5G15650.1 encodes for an enzyme (RGP2) highly expressed in actively dividing tissues such as pollen and is thought to play an important role in cell wall formation. Another gene, AT1G42470.1, encodes for an ortholog putative Niemann-Pick C1 protein involved in the steroid metabolic pathway and important for the formation of viable pollen.

The two genes, AT2G31220.1 and AT4G30520.1, encode respectively for a transcription factor (bHLH010) and a receptor like kinase both involved in the process related to the development of the anther. The two remaining genes, AT3G51550.1 and AT1G63530.1, encodes respectively (RLK) a receptor like kinase important for the interaction between ovule and pollen tube and then important for the fertilization process. The other gene, AT1G63530.1, encodes for a hypothetical protein playing a role in the interaction between the growing pollen tube and style; interaction that would affect the right growth of the pollen tube and consequently the fertilization of the ovule. From the perspective of the function of the gene enclosed in the second cluster, it seemed that it would represent all the processes going from anther and pollen formation to the early stages of flower fertilization when during the pollen tube protrusion into the style, specific molecular processes take place to lead the growth of the pollen tube in the right direction. All the six genes were downregulated in the mostly ‘crumbly’ samples suggesting that the under expression of these genes, controlling or more generally involved in anther and pollen formation as well as in the interaction pollen tube style, might contribute to the formation and growth of ‘crumbly’ like misshapen fruit in red raspberry.

3.3.2 Heatmap of microarray probes matching genes with GO terms related to flower development

The microarray analysis highlighted 12 probes, significantly differently expressed between the mostly and the never 'crumbly' phenotypes, with specific gene ontology (GO) terms related to flower development. These twelve probes were imported into GeneSpring software and following the procedure described in section 3.2.2, a tree cluster heatmap was created (Figure 3.4)

Again, as for the genes related to pollen, for those linked to flower development, the data of interest were primarily those at the open flower stage and secondarily of the closed bud stage. Again, the green berry stage was excluded because when the drupelets were already set, the study of the expression levels for the genes controlling any function related to flower development was less relevant. Two interesting clusters of probes, one on the top and the other on the bottom of the tree cluster heatmap were identified and showed in Figure 3.4; the probes reported on the right side of the heatmaps indicate the position of the two clusters.

The full list of the 12 probes selected to produce the tree cluster heatmap of Figure 3.4 was reported on Table 3.6. The seven probes located inside the two clusters highlighted in the heatmap of Figure 3.4 were reported in Table 3.7; for each probe were reported too, the gene ontology annotation and its code together with gene ID of the *Arabidopsis thaliana* genes ortholog of those from *Rubus idaeus* that were matched by the microarray probes.

In Tables 3.7 and 3.8 the *Arabidopsis thaliana* gene IDs for the 7 ortholog genes, of those *Rubus idaeus* matched by the 'crumbly' microarray experiment probes and included in the two highlighted cluster of potential interesting genes for the 'crumbly' fruit having gene ontology annotations related to flower development were reported. For each probe, in Tables 3.7 and 3.8, the predicted means with their standard error from the analysis of variance (ANOVA) with a stage (i.e. closed bud, open flower and green berry) per phenotype (i.e. mostly and never 'crumbly') interaction to show whether for each stage, the difference in the expression level of each probe was statistically significative at 95% confidence levels were reported.

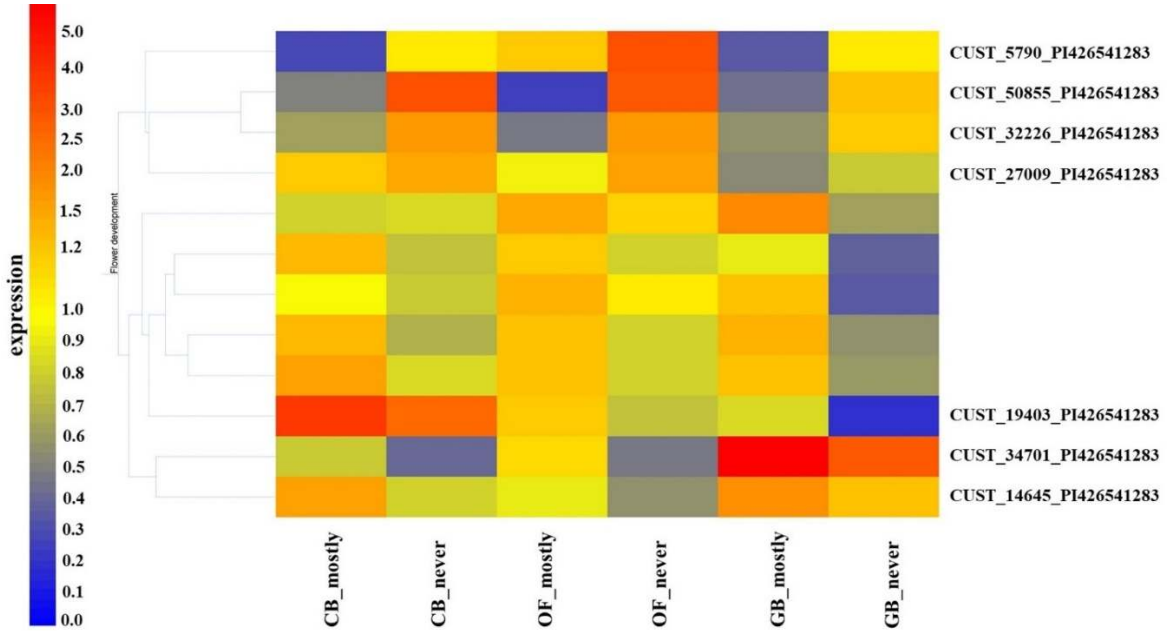


Figure 3.4: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to flower development.

Cluster tree heatmaps for the twelve microarray probes differently expressed for what concerned the stage*phenotype interaction, with differences being significative at 95% confidence levels. Three stages tested (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**)) and two phenotypes (i.e. mostly and never 'crumbly'). The twelve probes matched *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to flower development. Two main clusters of probes with difference in the expression level, as highlighted by the colour shades of the heatmap, were selected and the corresponding probes were reported on the right side of the heatmap. With respect to the first two stages (i.e. closed bud and open flower), the four probes within the first highlighted cluster were downregulated in the mostly 'crumbly' phenotype compared to the never one while those three within the second highlighted cluster were upregulated in the mostly 'crumbly' phenotype and upregulated in the never 'crumbly'. High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented.

	Microarray probe	Gene Ontology	GO code	*A. thaliana* gene ID
1st cluster	CUST_5790_PI426541283	flower development	GO:0009908	AT5G05660.1
	CUST_50855_PI426541283	flower development	GO:0009908	AT1G25540.2
	CUST_32226_PI426541283	embryo sac development	GO:0009553	AT3G12280.1
		double fertilization forming a zygote and endosperm	GO:0009567	
		gametophyte development	GO:0048229	
		endosperm development	GO:0009960	
	CUST_27009_PI426541283	double fertilization forming a zygote and endosperm	GO:0009567	AT4G12620.1
	CUST_55276_PI426541283	stamen development	GO:0048443	AT4G12620.1
		petal development	GO:0048441	
	CUST_52339_PI426541283	negative regulation of flower development	GO:0009910	AT1G79730.1
	CUST_34459_PI426541283	plant ovule development	GO:0048481	AT3G61120.1
	CUST_51962_PI426541283	flower development	GO:0009908	AT5G14530.1
	CUST_36819_PI426541283	flower morphogenesis	GO:0048439	AT5G55300.1
2nd cluster	CUST_19403_PI426541283	flower development	GO:0009908	AT1G69120.1
		floral meristem determinacy	GO:0010582	
	CUST_34701_PI426541283	negative regulation of flower development	GO:0009910	AT5G16260.1
	CUST_14645_PI426541283	regulation of flower development	GO:0009909	AT1G80940.1

Table 3.6: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to flower development.

List of the twelve microarray probes, significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significative at 95% confidence levels. The gene ID of the A. thaliana gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to flower development. The first column of the table indicates the probes, belonging to the two main clusters highlighted on the tree cluster heat map of Figure 3.4.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009908 (flower development)	AT5G05660.1	mostly	-0.56	0.078	-0.468	0.0901	CUST_5790_PI426541283
		never	0.009	0.509	0.011		
GO:0009908 (flower development)	AT1G25540.2	mostly	-0.301	-0.6	-0.365	0.1372	CUST_50855_PI426541283
		never	0.533	0.508	0.112		
GO:0009960 (endosperm development)	AT3G12280.1	mostly	-0.204	-0.325	-0.244	0.0701	CUST_32226_PI426541283
		never	0.247	0.259	0.076		
	AT4G12620.1	mostly	0.092	-0.026	-0.272	0.0417	CUST_27009_PI426541283
		never	0.195	0.233	-0.106		

[a]2 degrees of freedom and 4 replicates
[b]GO:0009567 (double fertilization forming a zygote and endosperm)

Table 3.7: ANOVA table of means for all the probes within the first cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to flower development.

ANOVA means of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the four probes within the first cluster and for the three probes within the second cluster, both highlighted in the flower development genes heatmap. For each probe the predicted means from the analysis of variance and the standard error were reported to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe the Arabidopsis thaliana gene ID and its gene ontology annotations, related to flower development, were reported.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009908 (flower development) [b]GO:0010582 (floral meristem determ.)	AT1G69120.1	mostly	0.613	0.088	-0.064	0.0504	CUST_19403_PI426541283
		never	0.437	-0.127	-0.7		
[c]GO:0009910 (negative regulation of flower dev.)	AT5G16260.1	mostly	-0.098	0.047	1.105	0.0985	CUST_34701_PI426541283
		never	-0.377	-0.331	0.489		
GO:0009909 (regulation of flower development)	AT1G80940.1	mostly	0.215	-0.04	0.289	0.0439	CUST_14645_PI426541283
		never	-0.077	-0.235	0.096		

[a]2 degrees of freedom and 4 replicates
[b]GO:0010582 (floral meristem determinacy)
[c]GO:0009910 (negative regulation of flower development)

Table 3.8: ANOVA table of means for all the probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to flower development.

ANOVA means of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the four probes within the first cluster and for the three probes within the second cluster, both highlighted in the flower development genes heatmap. For each probe the predicted means from the analysis of variance and the standard error were reported to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe the Arabidopsis thaliana gene ID and its gene ontology annotations, related to flower development, were reported.

For the four *Arabidopsis thaliana* genes, that matched probes within the first cluster of the heatmap (Figure 3.4), five different gene ontology terms related to flower development were found and the majority of them (Table 3.7) indicated a potential effect of these genes in controlling/regulating part of the processes linked with flower development; embryo sac formation and the processes immediately after the fertilization (i.e. zygote and endosperm formation). It would seem that the highlighted cluster, containing four genes, could represent the processes involved in flower formation and in particular of ovary (i.e. embryo sac), together with those of the double fertilization when zygote and endosperm are formed as a consequence of the fusion of the two sperm nuclei with egg cell and secondary nucleus respectively.

The function of the four genes identified in the first cluster of the heatmap in Figure 3.4 were:

- AT5G05660.1 – encodes NFXL2, an homologous of the mammalian transcription factor zinc finger transcription factor NF-X1, located in the nucleus and expressed in carpels (source TAIR https://ui.arabidopsis.org/) The gene seems to be involved in pollen germination and pollen tube growth (Wang et al., 2008).
- AT1G25540.2 – encodes MED25 a nuclear protein that acts in a phytochrome pathway (especially the phytochrome B (phyB) and downstream of it) and induces flowering in response to suboptimal light conditions. It is involved in jasmonic acid mediated signalling pathway, positive regulation of defense response and positive regulation of flower development. It is expressed in carpel tissue (source TAIR https://ui.arabidopsis.org/).
- AT3G12280.1 – encodes a retinoblastoma homologue RETINOBLASTOMA-RELATED protein (RBR or RBR1). RBR controls nuclear proliferation in the female gametophyte. It is also required for correct differentiation of male gametophytic cell types. It is involved in double fertilization forming a zygote and endosperm, embryo sac development, gametophyte development and generative cell differentiation. It is expressed in carpels (source TAIR https://ui.arabidopsis.org/).
- AT4G12620.1 – encodes the Origin Recognition Complex subunit 1b. Involved in the initiation of DNA replication UNE13, Unfertilised Embryo Sac 13; mutants showed unfertilized ovules but normal pollen tube attraction (Pagnussat et al., 2005).

An overview on the molecular functions performed by the four genes included in the first cluster, highlighted in Figure 3.4 showed that two genes, (i.e. AT3G12280.1 and AT4G12620.1) would seem to be involved and contributing to the fertilization of the

ovules. The gene AT5G05660.1 encoding a transcription factor (NFXL2) seems to contribute to pollen germination and pollen tube growth while the last gene AT1G25540.2 whose gene product (MED25) is a nuclear protein interacting downstream with the phytochrome B pathway seems to play a part in the positive regulation of flower development. Interestingly although all the four genes have gene ontology annotation related to flower development, three of them are potentially related to pollen functions such as pollen germination, pollen tube growth and pollen double fertilization of the ovule and only one gene product, MED25 encoded by AT1G25540.2, seems to be involved in flower development. From the perspective of the gene functions, this first cluster would represent the processes or part of them involved in the early steps of flower fertilization, starting from pollen germination on the stigma, continuing with pollen tube growth inside the style and finishing with the double fertilization of the ovules and central cell by the two sperm nuclei.

The four *Arabidopsis thaliana* genes, orthologs of those seven *Rubus idaeus* ones, matched by the microarray probes, were all downregulated in the mostly 'crumbly' phenotype compared to the never 'crumbly' one. This would suggest that the under-expression of these four genes might contribute to the formation of 'crumbly' like misshapen fruit in red raspberry.

The second cluster of probes, highlighted at the bottom of the tree cluster heatmap of Figure 3.4, contained three genes. Their gene ontology annotations were quite similar; three of them refers to flower development and its regulation while the remaining gene ontology terms references to 'floral meristem determinacy'. This second cluster would then represent the processes related to, in general, the formation of the flower.

The function of the three genes identified in the second cluster of the heatmap in Figure 3.4 were:

- AT1G69120.1 – encodes for a transcription factor APETALA1 which specifies floral meristem and sepal identity. AP1 is required for the transcriptional activation of AGAMOUS. APETALA1 interacts with LEAFY and binds the promoter to regulates the expression of flowering time genes such as: SVP, SOC1 and AGL24. (source TAIR https://ui.arabidopsis.org/).
- AT5G16260.1 - encodes an RNA binding protein ELF9 (EARLY FLOWERING9) that plays a role in the flower induction process. Flower induction is regulated by two different antagonistic pathways, floral promotion (depend on photoperiod) and floral

repression (depend on vernalization). The FLOWERING LOCUS T (FT) protein, activated by COSTANS (CO) stimulates floral transition and this function is counteracted by the floral repressive activity of the FLOWERING LOCUS C (FLC). These two pathways, the one stimulating and the other repressing floral transition converge to a point where other integrators are involved and within these the SUPPREESSOR of OVEREXPRESSION of CO1 (SCO1) whose function is to suppress CO1 (FT activator) and then repress floral transition. ELF9 is a RNA binding protein that targeting SOC1 transcript play a role in stimulating floral induction (Song et al., 2009).

- AT1G80940.1 - Snf1 kinase interactor-like protein is involved in regulation of flower development and expressed in flowers (source TAIR https://ui.arabidopsis.org/).

An overview on the molecular functions performed by these three genes, included in the second cluster, highlighted in Figure 3.4 showed that AT1G69120.1, encoding a transcription factor (APETALA1), involved in the activation of AGAMOUS, plays an indirect role in stimulating floral transition. In fact, AGAMOUS targets and represses WUSCHEL gene that is responsible for the maintenance of the vegetative meristem (Yamaguchi et al., 2017). The other gene AT5G16260.1 encodes ELF9, an RNA binding protein that targeting SCO1, the repressor of COSTANS (CO1). CO1 is the activator of the FLOWERING LOCUS T (FT) that in turn stimulates floral transition; ELF9 indirectly play a role in the processes regulating the transition from apical meristem to flower meristem. The last gene, AT1G80940.1, encodes for a kinase interactor-like protein (Snf1) involved in regulation of flower development. The three genes belonging to the second highlighted cluster of the 12 probes matching genes with gene ontology annotation related to flower development seem to represent part of the molecular processes that lead the transition from vegetative to floral meristem.

3.3.3 Heatmap of microarray probes matching genes with GO terms related to hormones

In total fifty-six probes matched *Rubus idaeus* genes whose *Arabidopis thaliana* counterparts had gene ontology annotations related to hormones. These fifty-six probes were imported into GeneSpring software and following the procedure described in section 3.2.2, a tree cluster heatmap, specific for these 56 probes, was created (Figure 3.5).

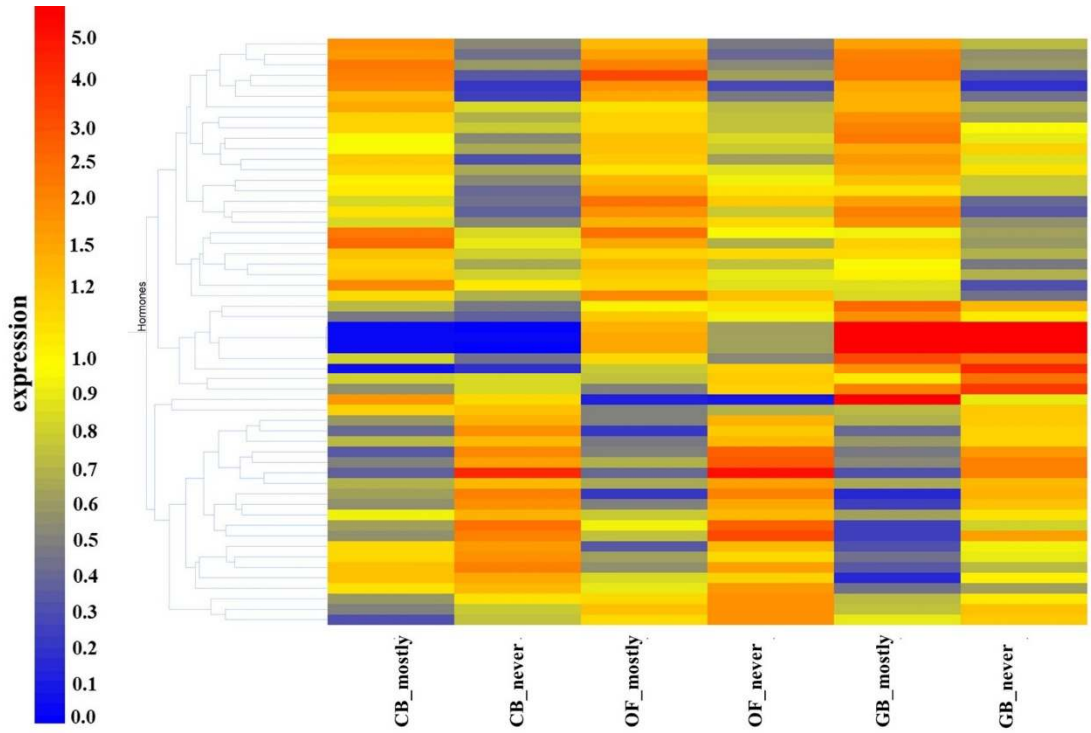


Figure 3.5: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to hormones.

Cluster tree heatmaps for the fifty-six microarray probes differently expressed for what concerned the stage*phenotype interaction with differences being significative at 95% confidence levels. Three stages tested (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**)) and two phenotypes (i.e. mostly and never 'crumbly'). The fifty-six probes matched *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to hormones. High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented.

The GO annotations related to hormones belonged to six groups (i.e. response, regulation, biosynthesis, signalling pathway, homestasis and metabolism). The expression levels of these probes were significantly different between the two different phenotype (i.e. mostly and never 'crumbly') with difference being significative at 95% confidence levels. Twelve of these probes, matching genes related to hormones, were located inside quantitative trait loci (QTLs) associated with 'crumbly' fruit that were mapped on linkage groups one and three (see chapter four for more details). In total fifty-three different GO terms, linked to hormones, were identified and the full list of genes with associated gene ontology annotations related to hormones.

The fifty-three different GO annotations linked with hormones were quite diverse and to facilitate interpretation of the data analysis, they were divided in three groups. The first group contained twenty-seven genes with GO annotation 'response to hormones', the second one with sixteen genes with GO term 'hormones biosynthesis' and the third one

was a miscellaneous group, called 'hormones other', containing the rest eighteen genes with any other GO terms related to hormones rather than the two previously mentioned. In some cases, because the same gene could have had more than one GO annotation, genes were found in more than one group. A tree cluster heatmap for each of the three group of GO terms (i.e. 'response to hormone', hormone biosynthesis' and 'hormones other') was created and each group was discussed separately.

3.3.3.1 Heatmap of microarray probes matching genes with GO annotations related to 'response to hormones'

The microarray analysis highlighted 27 probes, significantly differently expressed between the mostly and the never 'crumbly' phenotypes and with specific GO terms related to 'response to hormones'. These twenty-seven probes were imported into GeneSpring software and following the procedure described previously (see section 3.2.3), a tree cluster heatmap, specific for these 27 probes, was created (Figure 3.6).

The full list of the 17 probes highlighted in the three clusters of the tree cluster heatmap of Figure 3.6 was reported in Tables 3.9, for the first two clusters, and Table 3.10 for the third and last cluster of 7 probes at the bottom of the heatmap of Figure 3.6. All the remaining probes, those outside the three clusters, were listed in Table A.3.3 (see appendix). In Table 3.9 and 3.10, for each probe the *Arabidopsis thaliana* gene was reported the gene ontology term with corresponding code and the linkage group in case a gene was mapped in one of the three 'crumbly' QTLs located on linkage groups 1 and 3 (see chapter four for details).

In Tables 3.11, 3.12 and 3.13 the *Arabidopsis thaliana* gene IDs for the 17 ortholog genes, of those *Rubus idaeus*' one, matched by the 'crumbly' microarray experiment probes were reported. These seventeen probes were included in the three highlighted clusters in the heatmap of Figure 3.6 and contained potential interesting genes for the 'crumbly' fruit whose gene ontology annotations were related to 'response to hormone'.

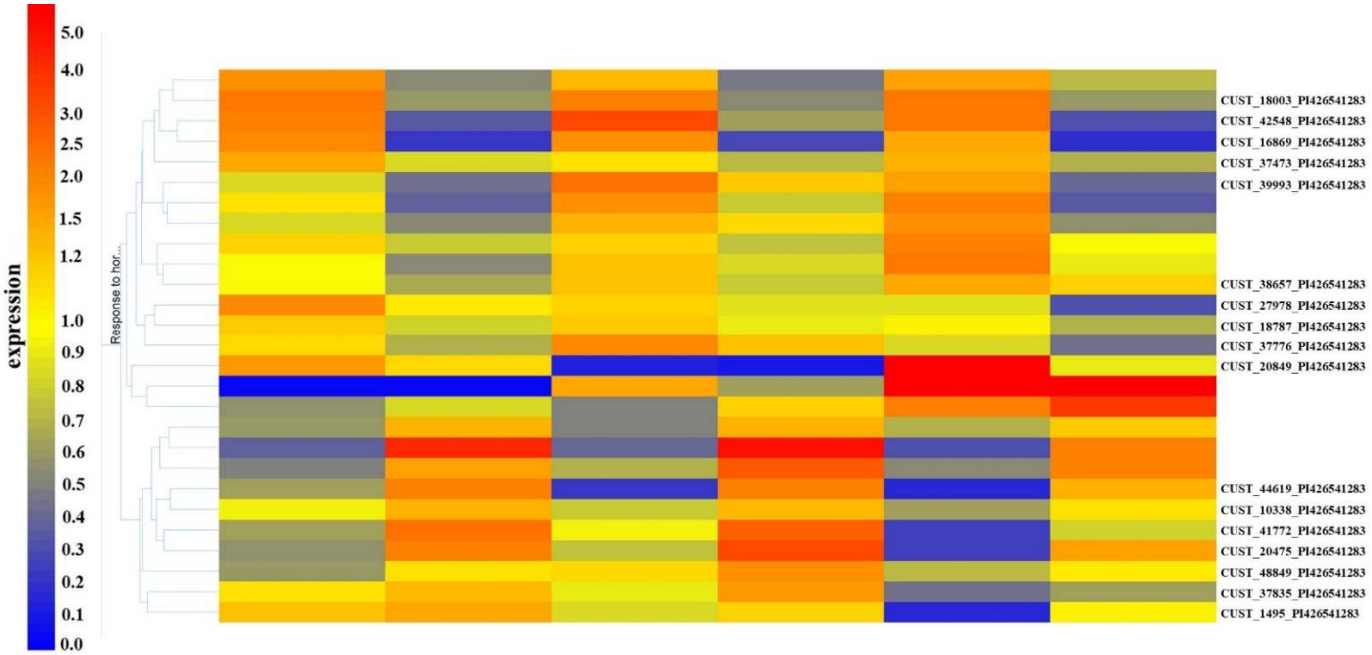


Figure 3.6: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to 'response to hormone'.

Cluster tree heatmaps for the twenty-seven microarray probes differently expressed between for what concerned the stage*phenotype interaction with differences being significative at 95% confidence levels. Three stages tested (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**)) and two phenotypes (i.e. mostly and never 'crumbly'). The twenty-seven probes matched *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to 'response to hormone'. Three main clusters of probes with evident difference in the expression levels, as highlighted by the colour shades of the heatmap, were selected and the corresponding probes were reported on the right side of the heatmap. The five probes within the first highlighted cluster, and those five within the second cluster, were upregulated in the mostly 'crumbly' phenotype compared to the never one. The last seven probes of the heatmap, those within the third cluster highlighted on the heatmap were downregulated in the mostly 'crumbly' phenotype and upregulated in the never 'crumbly'. High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented on the left side of the heatmap.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	linkage group
1st cluster	CUST_18003_PI426541283	response to abscisic acid (ABA)	GO:0009737	AT4G24520.1	
	CUST_42548_PI426541283	response to abscisic acid (ABA)	GO:0009737	AT2G33590.1	
	CUST_16869_PI426541283	response to abscisic acid ABA	GO:0009737	AT2G17840.1	
	CUST_37473_PI426541283	response to abscisic acid (ABA)	GO:0009737	AT5G51300.3	
	CUST_39993_PI426541283	response to abscisic acid (ABA)	GO:0009737	AT1G77120.1	
2nd cluster	CUST_38657_PI426541283	response to jasmonic acid (JA)	GO:0009753	AT2G46410.1	[a]LG3_old
		response to salicylic acid (SA)	GO:0009751		
	CUST_27978_PI426541283	response to salicylic acid (SA)	GO:0009751	AT1G15780.1	
	CUST_18787_PI426541283	response to Auxin (AUX)	GO:0009733	AT2G041600.1	[a]LG3_old
	CUST_37776_PI426541283	response to abscisic acid (ABA)	GO:0009737	AT5G25610.1	
	CUST_20489_PI426541283	response to cytokinin (CK)	GO:0009735	AT1G11910.1	[a]LG3_old
	[a]LG3_old – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3				

Table 3.9: Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'response to hormone'.

List of the ten microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the A. thaliana gene corresponding to the R. idaeus gene matched by the probe was reported together with the ontology term and its code; obviously related to 'response to hormone'. Moreover, for those probes matching genes mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), was reported the linkage group in which the QTL is located. In the first column of the table were indicated the clusters of the heatmap to which the probes belong.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	linkage group
3rd cluster	CUST_40382_PI426541283	response to cyclopentenone (OPDA)	GO:0010583	AT3G12110.1	[b]LG3_new
	CUST_10338_PI426541283	response to ABA	GO:0009737	AT3G48170.1	
	CUST_41772_PI426541283	response to cyclopentenone (OPDA)	GO:0010583	AT5G03050.1	
	CUST_20475_PI426541283	response to ABA	GO:0009737	AT1G55020.1	
		response to JA	GO:0009753		
	CUST_48849_PI426541283	response to SA	GO:0009751	AT5G49570.1	
	CUST_37835_PI426541283	response to Auxin	GO:0009733	AT1G30330.2	[a]LG3_old
	CUST_1495_PI426541283	response to ABA	GO:0009737	AT4G21410.1	

[a]LG3_old – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3
[b]LG3_new – mapped inside the new 'crumbly' QTL identified during this work and again on linkage group 3

Table 3.10: 'Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'response to hormone'.

List of the last seven microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the A. thaliana gene corresponding to the R. idaeus gene matched by the probe was reported together with the ontology term and its code; obviously related to 'response to hormone'. Moreover, for those probes matching genes mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), was reported the linkage group in which the QTL is located. In the first column of the table were indicated the clusters of the heatmap to which the probes belong.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009737 (response to ABA)	AT4G24520.1	mostly	0.379	0.33	0.362	0.0697	CUST_18003_PI426541283
		never	-0.217	-0.268	-0.219		
GO:0009737 (response to ABA)	AT2G33590.1	mostly	0.344	0.564	0.361	0.1049	CUST_42548_PI426541283
		never	-0.441	-0.19	-0.494		
GO:0009737 (response to ABA)	AT2G17840.1	mostly	0.324	0.272	0.189	0.0744	CUST_16869_PI426541283
		never	-0.671	-0.528	-0.748		
GO:0009737 (response to ABA)	AT5G51300.3	mostly	0.196	0.024	0.172	0.06	CUST_37473_PI426541283
		never	-0.066	-0.137	-0.151		
GO:0009737 (response to ABA)	AT1G77120.1	mostly	-0.062	0.396	0.219	0.0482	CUST_39993_PI426541283
		never	-0.343	0.095	-0.375		

[a]2 degrees of freedom and 4 replicates

Table 3.11: ANOVA table of means for all the five probes within the first cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'response to hormone.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the five probes within the first highlighted in the 'response to hormone' genes tree cluster heatmap (Figure 3.6). For each probe the predicted means from the analysis of variance and the standard error significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry) were reported. For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'response to hormone', were reported.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009753 (response to JA) GO:0009751 (response to SA)	AT2G46410.1	mostly	0.063	0.056	0.334	0.0598	CUST_38657_PI426541283[ł]
		never	-0.104	-0.114	-0.008		
GO:0009751 (response to SA)	AT1G15780.1	mostly	-0.01	0.115	0.378	0.0612	CUST_27978_PI426541283
		never	-0.257	-0.068	-0.03		
GO:0009751 (response to Auxin)	AT2G04160.1	mostly	-0.007	0.107	0.197	0.0632	CUST_18787_PI426541283[ł]
		never	-0.177	-0.093	0.055		
GO:0009737 (response to ABA)	AT5G25610.1	mostly	0.302	0.06[b]	-0.049	0.0643	CUST_37776_PI426541283
		never	0.014	-0.056[b]	-0.503		
GO:0009735 (response to CK)	AT1G11910.1	mostly	0.092	0.076	0.006	0.0373	CUST_20489_PI426541283[ł]
		never	-0.077	-0.032	-0.163		

[a]2 degrees of freedom and 4 replicates
[b]difference not statistically significant
[ł]probes mapped in the 'crumbly' QTL previously identified by Graham et al., (2015) on linkage group 3

Table 3.12: ANOVA table of means for all the five probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'response to hormone.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the five probes within the second highlighted in the 'response to hormone' genes tree cluster heatmap (Figure 3.6). For each probe the predicted means from the analysis of variance and the standard error significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry) were reported. For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'response to hormone', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located.

Gene Ontology (GO) term	A. thaliana gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0010583 (response to OPDA)	AT3G12110.1	mostly	-0.2	-0.669	-0.765	0.1335	CUST_44619_PI426541283[§]
		never	0.34	0.341	0.159		
GO:0009737 (response to ABA)	AT3G48170.1	mostly	-0.023	-0.099	-0.193	0.0509	CUST_10338_PI426541283
		never	0.159	0.144	0.025		
GO:0010583 (response to OPDA)	AT3G03050.1	mostly	-0.199	-0.02	-0.606	0.0797	CUST_41772_PI426541283
		never	0.401	0.455	-0.08		
GO:0009737 (response to ABA) GO:0009753 (response to JA)	AT1G55020.1	mostly	-0.232	-0.117	-0.584	0.0716	CUST_20475_PI426541283
		never	0.349	0.564	0.214		
GO:0009751 (response to SA)	AT5G49570.1	mostly	-0.227	0.051	-0.145	0.0508	CUST_48849_PI426541283
		never	0.021	0.287	0.016		
GO:0009751 (response to Auxin)	AT1G30330.2	mostly	0.021[b]	-0.037	-0.359	0.0646	CUST_37835_PI426541283[ł]
		never	0.145[b]	0.241	-0.201		
GO:0009737 (response to ABA)	AT4G21410.1	mostly	0.104[b]	-0.062[b]	-0.795	0.0765	CUST_1495_PI426541283
		never	0.193[b]	0.057[b]	0.006		

[a]2 degrees of freedom and 4 replicates
[b] difference not statistically significant
[§]probes mapped in the new 'crumbly' QTL identified here in this work on linkage group 3
[ł]probes mapped in the 'crumbly' QTL previously identified by Graham et al., (2015) on linkage group 3

Table 3.13: ANOVA table of means for all the seven probes within the third cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'response to hormone.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the five probes within the second highlighted in the 'response to hormone' genes tree cluster heatmap (Figure 3.6). For each probe the predicted means from the analysis of variance and the standard error significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry) were reported. For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'response to hormone', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located.

For each probe, in Tables 3.11, 3.12 and 3.13, the predicted means with their standard error from the analysis of variance (ANOVA) with a stage (i.e. closed bud, open flower and green berry) per phenotype (i.e. mostly and never 'crumbly') interaction to show whether for each stage, the difference in the expression level of each probe was statistically significative at 95% confidence levels were reported.

Of particular interest was the first cluster of probes identified on the tree cluster heatmap of Figure 3.6. All the five probes contained here matched *Rubus idaeus* genes whose *Arabidopsis thaliana* ortholog had gene ontology annotation related to response to ABA (abscisic acid). The differences were significant at 95% confidence levels for all the five probes across all three different development stages (i.e. closed bud, open flower and green berry) and the probes were all upregulated in the mostly 'crumbly' phenotype compared to the never one. Such finding would suggest the important role played by the genes activated by ABA in raspberry plants showing 'crumbly' fruits.

The function of the five genes identified in the first cluster of the heatmap in Figure 3.6 were:

- AT4G24520.1 - encodes the cytochrome450 (cyp450) reductase likely to be involved in phenylpropanoid metabolism (source TAIR https://www.arabidopsis.org/) Phenylpropanoids are involved in many physiological functions not only important for plant growth and development but for interaction with the external environment (Biała and Jasiński, 2018).
- AT2G33590.1 – encodes CRL1, a protein with homology to members of the dihydroflavonol-4-reductase (DFR) superfamily, operating in the phenylpropanoid pathway. The enzyme expression is induced by ABA and together with CRL2 is involved in the generation of vascular tissues (Ostergaard et al., 2001).
- AT2G17840.1 – encodes the Early-Responsive to Dehydration 7 (ERD7). This gene is upregulated by high light, drought, cold and salt stress and induced in response to ABA (source TAIR https://www.arabidopsis.org/)
- AT5G51300.3 – encodes the *Arabidopsis thaliana* SF1 (AtSF1) a nuclear localized splicing factor homolog that is involved in alternative splicing of some mRNAs (source TAIR https://www.arabidopsis.org/). A mutant allele of AtSF1 (AT5G51300) that contains a T–DNA insertion was shown to conferr pleiotropic

developmental defects, including early flowering and abnormal sensitivity to abscisic acid (Jang et al., 2014).

- AT1G77120.1 – encodes the *Arabidopsis thaliana* alcohol dehydrogenase 1 (ADH1). The enzyme catalyses the reduction of acetaldehyde using NADH as reductant and it requires zinc for its activity. The gene expression is induced by ABA (source TAIR https://www.arabidopsis.org/).

An overview on the molecular functions performed by these five genes included in the first cluster, highlighted on Figure 3.6 showed that three genes, (i.e. AT4G24520.1, AT2G33590.1 and AT1G77120.1), code for enzymes, the first two a reductase and the third an alcohol dehydrogenase. The two reductases both operate in the phenylpropanoid pathway and are involved in plant growth and development; for example, development of the vascular system as for the case of AT2G33590.1. The other two genes, AT2G17840.1 and AT5G51300.3, code respectively a protein ERD7, upregulated under specific abiotic stresses (i.e. drought, light and cold), and a component of the spliceosome. The product of these five genes were involved in generic physiological processes but all these genes are induced in response to ABA, highlighting the potential important role fulfilled by this specific hormone in the processes causing the formation of 'crumbly' like misshapen fruit. Such a finding is important and this is supported by the results from the metabolomic analysis (see chapter five for more details) where, statistically significant differences in the level of ABA were detected between artificial induced 'crumbly' fruit and normal raspberry fruit. The five *Arabidopsis thaliana* genes, orthologs of the five *Rubus idaeus* genes, matched by the microarray probes, were all upregulated in the mostly 'crumbly' phenotype compared to the never 'crumbly' one. This would suggest that overexpression of these five genes would require high level of ABA. The metabolomic analysis (see chapter 5), interestingly, showed significantly higher levels of ABA in the receptacle at the green berry stage, in artificially induced 'crumbly' fruit (see chapter 5 for more details). On the basis of these results, it can be concluded that a potential role played by ABA is contributing to the formation of 'crumbly' like misshapen fruit in red raspberry.

The second cluster of probes, identified on the tree cluster heatmap of Figure 3.6, contained five probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had five different gene ontology annotation related to 'response hormone' (see Table 3.9). In details these five ontology terms were: response to salicylic acid (SA),

jasmonic acid (JA), auxin (Aux), ABA and cytokinin (Ck), indicating that the second cluster would represent the molecular/physiological process activated in response to the hormonal stimuli exerted by all these five different hormones. It was important to notice that three of these genes (i.e. AT2G46410.1, AT2G04160.1 and AT1G11910.1) had microarray probes that mapped inside the 'crumbly' QTL (see chapter four for more details) identified by Graham et al. (2015) giving further support to their effect on the induction of 'crumbly' fruit of red raspberry. For all five probes belonging to the second highlighted cluster, the differences in the expression levels were significative at 95% confidence levels in all the three different development stages (i.e. closed bud, open flower and green berry) and the probes were all upregulated in the mostly 'crumbly' phenotype compared to the never one.

The function of the five genes identified in the second cluster of the heatmap in Figure 3.6 were:

- AT2G46410.1 – encodes CAPRICE (CPC) a R3-type MYB transcription factor trichome with a positive regulation effect of hair-cell differentiation. The transcription factor is preferentially transcribed in hairless cells (source TAIR - https://www.arabidopsis.org/).
- AT1G15780.1 – encodes NRB4, mediator of RNA polymerase II transcription subunit. The mediator is a complex required for the normal transcription of genes. NRB4 is strongly activated in response to salicylic acid (SA). (Canet et al., 2012) with NBR4 null mutants (nbr4), found that not only were the mutants smaller in size but they were sterile and the plants bore flowers with no stamens and with abnormal carpels. The authors postulated that the specific phenotypes of the nrb4 null plants are the consequence of lack of response to salicylic acid (SA); such hypothesis would support the postulated essential role of SA in normal plant development.
- AT2G04160.1 – encodes AIR3, a protein similar to the subtilisin-like serine protease which is believed to be active outside the plant cell (source TAIR - https://www.arabidopsis.org/).
- AT5G25610.1 – encodes the protein Responsive to Desiccation 22 (RD22), activated in response to ABA under abiotic stresses (i.e. drought, salinity and light) (source TAIR - https://www.arabidopsis.org/).

- AT1G11910.1 – encodes Aspartic Proteinase A1 (APA1), an enzyme involved in ABA-dependent response thats overexpression confers drought tolerance in Arabidopsis (Sebastian et al., 2020).

The analysis of the gene products for the five genes whose microarray probes were found inside the second cluster highlighted in the heatmap on Figure 3.6 showed some interesting results. Three out of five of these differently expressed probes were mapped inside the 'crumbly' QTL previously identified by Graham et al. (2015); a feature that supports their role at some stage of 'crumbly' fruit development (see Table 3.9). Of these three genes matched by probes mapped in the 'crumbly', AT2G46410.1 might be relevant. It encodes a R3-type MYB transcription factor (CPC) that is expressed in hairless cells. The gene seems to be important in trichome development and in fact *Arabidopsis thaliana* transgenic plants overexpressing CPC lacked trichomes on leaves, stems and sepals (Wada et al., 1997). The microarray probe matching AT2G46410.1 was upregulated in the mostly 'crumbly' plants with the highest difference in the expression levels at green berry stage. In raspberry the drupelets present abundant epidermal hairs (i.e. unicellular linear trichomes) at their base and side. The entanglement of these hairs provides the cohesion to the drupelets and the shape stability of the berry (Jennings, 1988). It cannot be excluded that the upregulation of CPC in the mostly 'crumbly' plants might cause the formation of drupelets with reduced number of trichomes so they cannot adhere perfectly to each other and cause the fruit to crumble.

An exception might be AT1G15780.1 whose gene product, the mediator of RNA polymerase II transcription subunit (NRB4), involved in an important biological function (DNA transcription) and strongly activated in response to salicylic acid, would seem to be important for normal growth of plants and for fertility. On the basis of these considerations, the role played by this gene (AT1G15780.1), in contributing to the formation of 'crumbly' fruit, cannot be excluded.

All the probes within the second identified cluster of the tree cluster heatmap of Figure 3.6 were upregulated in the mostly 'crumbly' phenotype compared to the never 'crumbly'. The overexpression might contribute to cause the formation of 'crumbly' like misshapen fruit.

In the third highlighted cluster of probes of the heatmap (see Figure 3.6), the *Rubus idaeus* genes had *Arabidopsis thaliana* orthologs with five gene ontology annotations related to 'response to hormone' (see Table 3.10). The annotation 'response to ABA' was the most

abundant and present on three different genes confirming again the potential importance of this hormone and of the molecular/physiological function it could induce in the cells/plant for 'crumbly' fruit. The other class of hormones was the Jasmonates, and two genes had gene ontology annotation 'response to cyclopentanone' (OPDA) which is both a precursor of jasmonic acid (JA) and has growth regulating activity (Stintzi et al., 2001). One of the two genes with annotation, 'response to cyclopentanone' (OPDA), was mapped inside the new 'crumbly' QTL identified during this work (see chapter four for more details) giving further support to its potential involvement causing 'crumby' fruit in red raspberry. The other gene (AT1G30330.2) mapped to a previously identified 'crumbly' QTL Graham et al. (2015). For AT1G30330.2, the gene ontology annotation was 'response to Auxin' while for the last gene (AT5G49570.1) of this cluster the ontology term was 'response to SA'. This third cluster represents the role played by four classes of hormones (jasmonates, ABA, auxin and salicylic acid) whose effect could cause anomalous responses in the plants leading to the formation of misshapen 'crumbly' like fruits. For all these seven probes the differences in the expression levels were significative at 95% confidence levels in all the three different development stages (i.e. closed bud, open flower and green berry) and the probes were all downregulated in the mostly 'crumbly' phenotype compared to the never one.

The function of the seven genes identified in the third cluster of the heatmap in Figure 3.6 were:

- AT3G12110.1 – encodes an actin (ACT11) that is expressed predominantly during reproductive development. The *Arabidopsis thaliana* gene AT3G12110.1, encodes the reproductive actin ACT11. Reproductive actins are expected to play a part in the cytoskeleton rearrangement. ACT11 seems to play an important role in pollen germination and pollen tube elongation. In fact, loss of ACT11 reduces the amount of total actin in the cells with a consequent reduction/alteration of filamentous actin (F-actin) levels that negatively impact on both pollen germination and pollen tube growth elongation. Actin filaments are thought to form the molecular track, responsible for the transport of material indispensable for both membrane expansion and cell wall synthesis that allow the protrusion of the pollen tube (Chang and Huang, 2015). The raspberry gene corresponding to the *Arabidopsis thaliana* AT3G12110.1, was downregulated in the mostly 'crumbly' phenotype with difference being significant at 95% of confidence level in all the three stages analysed, but with the biggest difference (≈ 10 fold less compared to the never 'crumbly') in the open flower

stage. Moreover, the *Rubus idaeus*, equivalent of AT3G12110.1, mapped in the new 'crumbly' QTL on linkage group 3, recently identified during this work (for more details see chapter 4).

- AT3G48170.1 – encodes ALDH10A9, a protein that can function as a betaine aldehyde dehydrogenase in vitro. ALDH10A9 transcript levels rise in response to ABA (source TAIR - https://www.arabidopsis.org/)
- AT3G03050.1 – encodes a Cellulose Synthase-like D3 (CSLD3), a cellulose synthase like protein involved in the synthesis of non-cellulosic wall polysaccharide whose gene is downregulated in response to cyclopentanone (source TAIR - https://www.arabidopsis.org/).
- AT1G55020.1 – encodes a Lipoxygenase 1 (LOX1), a defense gene conferring resistance *Xanthomonas campestris* (source TAIR - https://www.arabidopsis.org/).
- AT5G49570.1 – encodes a protein with N-glycanase activity; the enzyme is Peptide-N-Glycanase 1 (PNG1). The protein is a component of the ERAD system; a multiple-step degradation process, located in the endoplasmic reticulum (ER) that removes all the misfolded new synthesized protein that have been previously quality checked by the endoplasmic reticulum-mediated protein quality control (ERQC) mechanism. ERQC/ERAD systems are involved in the regulation of important biochemical/physiological processes, such as abiotic stress tolerance and plant defense (Liu and Li, 2014).
- AT1G30330.2 - encodes a member of the auxin response factor (ARF) family. It mediates auxin response via expression of auxin regulated genes. Specifically, the gene encodes ARF6, that has been shown to be responsible for the correct maturation of stamens and for fertility of the plant (Liu et al., 2018). It might be the case that the mostly 'crumbly' type plants behave like *Arabidopsis thaliana* arf6 mutants that showed delayed stamen development and reduced fecundity. The gene was mapped in the 'crumbly' QTL on linkage group 3 (see chapter 4 for more details) previously identified by Graham et al. (2015).
- AT4G21410.1 - encodes a cysteine-rich receptor-like protein kinase (CRK29) involved in protein phosphorylation and activated in response to ABA (source TAIR - https://www.arabidopsis.org/).

The analysis of the functions of the seven genes contained in the third cluster, highlighted at the bottom of the heatmap (see Figure 3.6), showed that four genes (i.e. AT3G48170.1,

AT3G03050.1, AT1G55020.1 and AT4G21410.1) encodes protein/enzymes with a genetic function which was difficult to interpret in light of their potential contribution to the 'crumbly' fruit. One gene, AT3G03050.1, encoding (CSLD3), a cellulose synthase like protein involved in the synthesis of non-cellulosic wall polysaccharide, preferentially expressed in carpels and pollen (source TAIR - https://www.arabidopsis.org/) and whose expression is downregulated in response to hormonal stimulus (OPDA). AT3G03050.1 might be considered a potential important candidate for 'crumbly' fruit because impairment in the synthesis of cell wall components can negatively affect cell elongation, especially for fast growing tissues like those in the pollen tube and then compromise the fertilization of the ovule. The importance of the molecular processes involved in cell wall formation have already been discussed in this chapter; in fact the three genes (i.e. AT3G01640.1 , AT1G63180.1 and AT2G13680.1) would seem to play a role in the formation of the cell wall of the pollen grains (see section 3.3.1 of this chapter). We cannot exclude a priori this could not be the case for gene AT3G03050.1, especially in light of the fact that its gene product (i.e. cellulose synthase like protein) is involved in the biosynthesis of cell wall components, and that for instance pollen and pollen tube are within the gene AT3G03050.1 specific site of expression (source TAIR - https://www.arabidopsis.org/).

The last two genes, AT3G12110.1 and AT1G30330.2, were candidates for processes/functions causing the formation and development of 'crumbly' like fruit. The first gene encodes (ACT11) a component of the machinery (cytoskeleton) that seems to control the transport of material indispensable for both membrane expansion and cell wall synthesis that allow the protrusion of the pollen tube. Anomalous assembly of the components of this this machinery (cytoskeleton), ACT11 within these, can negatively affect pollen germination and pollen tube growth compromising the fertilization process. The second gene AT1G30330.2, encoding a response transcription factor ARF9, seems to be involved in the maturation of the stamen and then of the correct formation of pollen. Both these genes were differentially expressed between the two phenotypes and both were downregulated in the mostly 'crumbly' phenotype. In addition, both genes were mapped inside the two 'crumbly' QTL on linkage group 3 (see chapter four for more details) further supporting a key role.

3.3.3.2 Heatmap of microarray probes matching genes with gene ontology related to 'hormone biosynthesis'

The microarray analysis highlighted 16 probes, significantly differentially expressed for what concerned the stage*phenotype interaction with specific GO terms related to 'hormone biosynthesis'. These sixteen probes were imported into GeneSpring software and following the procedure described in section 3.2.2, a tree cluster heatmap, specific for these 16 probes, was created (Figure 3.7). Two main cluster of probes were highlighted respectively on top and bottom of the heatmap and the corresponding probes within each cluster were reported on the right side of the heatmap (see Figure 3.7).

The full list of the twelve probes comprised in the two clusters highlighted at the top and bottom of the tree cluster heatmap of Figure 3.7 was reported on Table 3.14. The list of remaining five probes, those outside the two clusters, always of the tree cluster heatmap of Figure 3.7, was reported in Table A.3.4 in appendix. For each probe the *Arabidopsis thaliana* ortholog gene to the *Rubus idaeus*' one, matched by the microarray probe, was reported.

In Tables 3.15 and 3.16, the *Arabidopsis thaliana* gene IDs for the 12 *Arabidopsis thaliana* genes, whose ortholog *Rubus idaeus*' one, were matched by the 'crumbly' microarray experiment probes were reported. The table included only those probes contained in the two highlighted cluster of potential interesting genes for the 'crumbly' fruit that have gene ontology annotations related to 'hormone biosynthesis'. For each probe, in Tables 3.15 and 3.16, the predicted means with their standard error from the analysis of variance (ANOVA) with a stage (i.e. closed bud, open flower and green berry) per phenotype (i.e. mostly and never 'crumbly') interaction to show whether for each stage, the difference in the expression level of each probe was statistically significative at 95% confidence levels were reported.

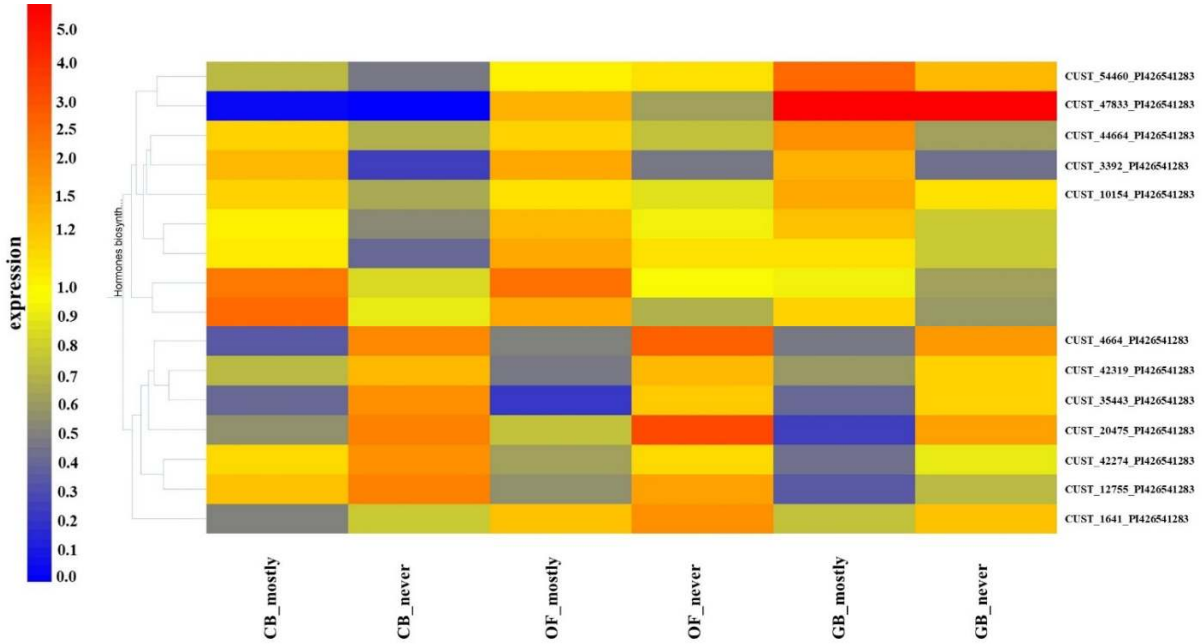


Figure 3.7: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to 'hormone biosynthesis'.

Cluster tree heatmaps for the sixteen microarray probes differently expressed for what concerned the stage*phenotype interaction with differences being significative at 95% confidence levels. Three stages tested (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**)) and two phenotypes (i.e. mostly and never 'crumbly'). The sixteen probes matched *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to 'hormone biosynthesis'. Two main clusters of probes with evident difference in the expression level, as highlighted by the colour shades of the heatmap, were selected and the corresponding probes were reported on the right side of the heatmap. The five probes, within the first highlighted cluster, were upregulated in the mostly 'crumbly' phenotype compared to the never one. The last seven probes of the heatmap, those within the second cluster highlighted on the heatmap were downregulated in the mostly 'crumbly' phenotype and upregulated in the never 'crumbly'. High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented on the left side of the heatmap.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	linkage group
1st cluster	CUST_54460_PI426541283	cytokinin biosynthetic process	GO:0009691	AT1G62360.1	[a]LG3_old
	CUST_47833_PI426541283	brassinosteroids biosynthetic process	GO:0016132	AT5G36140.1	
	CUST_44664_PI426541283	JA biosynthetic processes	GO:0009695	AT2G33150.1	
	CUST_3392_PI426541283	cinnamic acid biosynthetic process	GO:0009800	AT3G53260.1	
	CUST_10154_PI426541283	cytokinin dehydrogenase activity	GO:0019139	AT3G63440.1	[b]LG3_new
2nd cluster	CUST_4664_PI426541283	methyl indole-3-acetate esterase activity	GO:0080030	AT2G03550.1	
	CUST_42319_PI426541283	sterol biosynthetic process	GO:0016126	AT3G02590.1	
	CUST_35443_PI426541283	sterol biosynthetic process	GO:0016126	AT1G58440.1	
	CUST_20475_PI426541283	JA biosynthetic processes	GO:0009695	AT1G55020.1	
	CUST_42274_PI426541283	methyl indole-3-acetate esterase activity	GO:0080030	AT1G19190.1	
	CUST_12755_PI426541283	sterol biosynthetic process	GO:0016126	AT1G76490.1	
	CUST_1641_PI426541283	methyl indole-3-acetate esterase activity	GO:0080030	AT3G48690.1	
	[a]LG3_old - mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3 [b]LG3_new - mapped inside the new 'crumbly' QTL identified during this work and again on linkage group 3				

Table 3.14: 'Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'hormone biosynthesis'.

List of the twelve microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to 'hormone biosynthesis'. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table were indicated the clusters of the heatmap to which the probes belong.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009691 (cytokinin biosynthetic process)	AT1G62360.1	mostly	-0.14	0.007[b]	0.435	0.0454	CUST_54460_PI426541283[ł]
		never	-0.324	0.021[b]	0.133		
GO:0016132 [c](brassinos. biosynthetic process)	AT5G36140.1	mostly	-1.454	0.171	2.117	0.0751	CUST_47833_PI426541283
		never	-1.845	-0.188	1.903		
GO:0009695 (JA biosynthetic process)	AT2G33150.1	mostly	0.07	0.068	0.294	0.0629	CUST_44664_PI426541283
		never	-0.156	-0.122	-0.19		
GO:0009800 [d](CA biosynthetic process)	AT3G53260.1	mostly	0.126	0.178	0.172	0.0538	CUST_3392_PI426541283
		never	-0.569	-0.309	-0.362		
GO:0009691 [e](CK dehydrogenase activity)	AT3G63440.1	mostly	0.055	0.03[b]	0.205	0.0485	CUST_10154_PI426541283[§]
		never	-0.185	-0.044[b]	0.027		

[a]2 degrees of freedom and 4 replicates
[b]difference not statistically significant
[c]brassinosteroids biosynthetic process
[d]cinnamic acid biosynthetic process
[e]cytokinin dehydrogenase activity
[ł]probes mapped inside the 'crumbly' QTL previously identified by Graham et al., (2015) on linkage group 3
[§]probes mapped inside the new 'crumbly' QTL identified here in this work on linkage group 3

Table 3.15: ANOVA table of means for all the probes within the first cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'hormone biosynthesis.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the five probes within the first cluster highlighted in the 'hormone biosynthesis' genes tree cluster heatmap (Figure 3.7). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'hormone biosynthesis', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL is located**.**

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0080030 [b](MeIAA esterase activity)	AT2G03550.1	mostly	-0.452	-0.306	-0.308	0.0669	CUST_4664_PI426541283
		never	0.299	0.474	0.261		
GO:0016126 (sterol biosynthetic process)	AT3G02590.1	mostly	-0.143	-0.312	-0.215	0.0609	CUST_42319_PI426541283
		never	0.13	0.143	0.068		
GO:0016126 (sterol biosynthetic process)	AT1G58440.1	mostly	-0.394	-0.647	-0.373	0.1277	CUST_53443_PI426541283
		never	0.276	0.096	0.06		
GO:0016126 (sterol biosynthetic process)	AT1G55020.1	mostly	-0.232	-0.117	-0.584	0.0716	CUST_20475_PI426541283
		never	0.349	0.564	0.214		
GO:0080030 [b](MeIAA esterase activity)	AT1G19190.1	mostly	0.036	-0.199	-0.356	0.053	CUST_42274_PI426541283
		never	0.279	0.036	-0.035		
GO:0016126 (sterol biosynthetic process)	AT1G76490.1	mostly	0.11	-0.248	-0.464	0.0752	CUST_12755_PI426541283
		never	0.338	0.216	-0.129		
GO:0080030 [b](MeIAA esterase activity)	AT3G48690.1	mostly	-0.3	0.107	-0.125	0.0493	CUST_1641_PI426541283
		never	-0.105	0.294	0.117		

[a]2 degrees of freedom and 4 replicates
[b] methyl indole-3-acetate esterase activity

Table 3.16: ANOVA table of means for all the probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'hormone biosynthesis.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction, for the seven probes within the second cluster highlighted in the 'hormone biosynthesis' genes tree cluster heatmap (Figure 3.7). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'hormone biosynthesis', were reported.

In the first highlighted cluster of probes of the heatmap (see Figure 3.7), the *Rubus idaeus* genes that matched *Arabidopsis thaliana* orthologs with five different gene ontology annotations related to 'hormone biosynthesis' (see Table 3.14) were reported. Interestingly, the first and last probes of this cluster both mapped inside the two 'crumbly' QTLs (see Table 3.14) and they both had a gene ontology annotation related to cytokinins (Cks). However, while the first gene concerns Cks biosynthesis the other one, is involved in their degradation. The remaining three probes of this cluster were all related, again from the gene ontology term perspective, to the biosynthesis of, jasmonic acid, brassinosteroids and cinnamic acid; this last one being primarily an intermediate in the salicylic acid biosynthesis.

This first cluster could represent the role played by these three classes of hormones (i.e. cytokinin, jasmonic acid, cinnamic acid and brassinosteroids) whose biosynthesis could potentially activate molecular/physiological processes whose effects could lead to the formation of misshapen 'crumbly' like fruits. For all the five probes, belonging to the first highlighted cluster (Figure 3.7), the differences in the expression levels were significant at 95% confidence levels (Table 3.15) in all the three different development stages (i.e. closed bud, open flower and green berry); with the first probe, but only for the open flower stage, being the exception. All the probes were upregulated in the mostly 'crumbly' phenotype compared to the never one, suggesting that an increase in the biosynthesis of these hormones could play a role in causing 'crumbly' fruit in red raspberry.

The function of the five genes identified in the first cluster of the heatmap in Figure 3.7 were:

- AT1G62360.1 – encodes the Shoot Meristemless (STM), a Class 1 Knotted1-LIKE Homeobox (KNOXI) transcription factor important for the establishment and maintenance of the shoot apical meristem (SAM). Recent studies have shown how STM is involved in flower differentiation. STM mutants (stm) were shown to have impaired flowers with lower number of petals and stamens and no carpels. Overexpression of STM causes the formation of enlarged gynoecia consisting of carpeloid ovules (Roth et al., 2018). The gene was upregulated in the mostly 'crumbly' phenotype with the difference being significative at 95% confidence levels at closed bud stage (Table 3.15), when flower differentiation might take place, and not at open flower stage.

- AT5G36140.1 – encodes a member of the CYP716A subfamily of the cytochrome P450 monooxygenases with triterpene oxidizing activity, involved in sterol and triterpenoid metabolisms (source TAIR - https://www.arabidopsis.org/)
- AT2G33150.1 - encodes an organellar (peroxisome, glyoxysome) 3-ketoacyl-CoA thiolase (KAT2). The enzyme positively regulates ABA signalling in all the major ABA responses and since KAT2 was shown to be involved in the production of reactive oxygen species (ROS) it was suggested that the regulation of ABA signalling goes through the modulation of ROS homeostasis in the cells (Jiang et al., 2011). The presence of a gene involved in ABA signalling is particular interesting in the light of the metabolomic (hormones) analysis discussed in chapter five where ABA, measured in the receptacle, was the only hormone, out of the eighteen detected/quantified, whose amounts were statistically different between normal fruit and artificially induced 'crumbly' one (see chapter five for more details). The hormones signalling is important for the proposed hypothetical hormonal crosstalk, between receptacle and fertilized ovaries. This interplay, with the receptacle acting as central hub, should regulate and synchronise the growth of all the fertilised ovaries favouring the homogeneous growth of the fruit and potentially avoiding the formation of misshapen fruit such as those found in 'crumbly' fruit affected plants.
- AT3G53260.1 – encodes a Phenylalanine Ammonia-Lyase 2 (PAL2) and enzyme involved in the biosynthetic process of cinnamic acid. (source TAIR - https://www.arabidopsis.org/)
- AT3G63440.1 - encodes Cytokinin oxidase 6 (CKX6 previously called CKX7) that catalyses the irreversible inactivation of CKs, transforming them to adenine and a corresponding side chain-derived aldehyde. In fruit CKX7 degrades CKs that have negative effect upon fruit elongation (Di Marzo et al., 2020).

The analysis of the five genes within the first cluster highlighted in the tree cluster heatmap of Figure 3.7 are currently difficult to interpret from a 'crumbly' fruit perspective. For two genes (i.e. AT5G36140.1 and AT3G53260.1) whose product are enzymes involved in the biosynthesis of hormones precursors this is easier to relate. The gene AT2G33150.1 proved to be very interesting because its product (KAT2) is an enzyme that positively regulates ABA signalling. Hormonal signalling was one of the basic ideas behind part of this work; a hormonal crosstalk between the flower first and then of the fruit (i.e. receptacle and fertilized ovaries), leads to the correct growth and development of a complex fruit such as raspberry. The two last genes (i.e. AT1G62360.1

and AT3G63440.1) were very interesting functionally and because they were mapped inside the two 'crumbly' QTL on linkage group 3, which gives further remarks to their potential involvement in processes relating to 'crumbly' fruit. The first, AT1G62360.1, encodes a transcription factor (STM) which play important roles in the regulation of flower differentiation while the other gene, AT3G63440.1, encodes (CKX7) a cytokinin dehydrogenase, an enzyme, responsible for the degradation of the cytokinins which negatively affect fruit elongation. All the genes were upregulated in the mostly 'crumbly' phenotype and the differences in the expression levels were statistically significative at any of the three stages tested (Table 3.15). In conclusion the first cluster of probes, referring to the probes matching genes with gene ontology terms related to 'hormone biosynthesis' could represent anomalous processes in flower and fruit development that could be responsible of 'crumbly' fruit.

The second cluster of microarray probes differentially expressed between the two phenotypes (i.e. mostly and neve 'crumbly), with difference being significative at 95% confidence level (Table 3.16) contains seven probes and their *Rubus idaeus*' genes had *Arabidopsis thaliana* orthologs with gene ontology annotation related to 'hormone biosynthesis. Only three gene ontology terms were represented in this cluster, 'methyl indole-3-acetate esterase activity', 'sterol biosynthetic process' and 'JA biosynthetic process' (see Table 3.14 for details). Particularly interesting, in respect to 'crumbly' fruit and specifically in connection with the hormonal crosstalk hypothesis behind this work, was the ontology term 'methyl indole-3-acetate esterase activity'. The enzyme catalyses the conversion of methyl indole-3-acetate (MeIAA) to indole-3-acetate (IAA) and with such reaction, an acid compound (IAA) is converted in a nonpolar compound (MeIAA). Nonpolar molecules should be able to cross membranes easily; MeIAA could represent a transport form of IAA that is moved from a site of production to a site of action where it is then transformed into active IAA by the methylesterases (Yang et al., 2008).

The function of the seven genes identified in the second cluster of the heatmap in Figure 3.7 were:

- AT2G03550.1 – encodes alpha/beta-Hydrolases superfamily protein with methyl indole-3-acetate esterase activity expressed during flowering stage (source TAIR - https://www.arabidopsis.org/)

- AT3G02590.1 – encodes a fatty acid hydroxylase superfamily protein involved in the sterol biosynthetic processes and expressed during flower stage (source TAIR - https://www.arabidopsis.org/)
- AT1G58440.1 – encodes the Squalene Epoxidase 1 (SQE1) which converts squalene into 2,3-oxidosqualene the precursor of plant sterols. These compounds are very important since they are the biosynthetic precursors of steroid hormones in animals, insects and plants. Evidences suggests the role of sterols as signalling molecules (Pose and Botella, 2009).
- AT1G55020.1 - encodes a Lipoxygenase 1 (LOX1), a defense gene conferring resistance *Xanthomonas campestris* (source TAIR - https://www.arabidopsis.org/).
- AT1G19190.1 - encodes alpha/beta-Hydrolases superfamily protein with methyl indole-3-acetate esterase activity expressed during flowering stage (source TAIR - https://www.arabidopsis.org/)
- AT1G76490.1 - encodes a 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR1) involved in sterol biosynthesis during male gametogenesis. Impaired expression of this gene causes anomalous pollen formation due to defects in the membrane systems that in turn depend on sterols whose biosynthesis is regulated by HMG genes and HMG1 within these (Suzuki et al., 2009).
- AT3G48690.1 – encodes CarboxyEsterase 12 (CXE12) with methyl indole-3-acetate esterase activity which is expressed during flowering stage (source TAIR - https://www.arabidopsis.org/).

All seven genes highlighted in the second cluster of the heatmap of Figure 3.7 were downregulated in the mostly 'crumbly' phenotype compared to the never one. This would suggest that the under-expression of these gene might cause alteration in molecular/physiological processes whose effect could be the formation of 'crumbly' like misshapen fruit. Three of these seven genes (AT2G03550.1, AT1G19190.1 and AT3G48690.1) encode different enzymes but all with methyl indole-3-acetate esterase activity that convert methyl indole-3-acetate (MeIAA) to indole-3-acetate (IAA). MeIAA being nonpolar can more easily move through membranes (Yang et al. 2008) and could represent a molecular signal used by plants to signal between different parts/tissues in order to drive certain stimuli in a targeted region. The downregulation of these three genes in the mostly 'crumbly' might suggest that in the plants showing 'crumbly' fruit such hypothetical hormonal crosstalk would be impaired and this would negatively affect fruit

development. Three genes (i.e. AT3G02590.1, AT1G58440.1 and AT1G76490.1) are involved in the sterol biosynthesis. Sterols that are very important not only as precursors of steroids but because they may have important signalling function. Such interpretation of one of the roles played by sterol is again interesting in light of the hypothesised molecular crosstalk between receptacle and fertilised ovaries that should lead the growth and development of the fruit in the right way. The gene AT1G76490.1, in connection with 'crumbly' fruit, was interesting since it is involved in the synthesis of sterol but is expressed particularly during male gametogenesis playing an important role in the formation of viable pollen. The second and last cluster of probes, for genes with gene ontology terms related to 'hormone biosynthesis' could represent a potential system for conveying molecular messages.

3.3.3.3 Heatmap of microarray probes matching genes with GO terms related to 'hormones other'

The microarray analysis highlighted 18 probes, significantly differentially expressed between the mostly and the never 'crumbly' phenotypes, with specific GO terms related to other hormone annotations other than those previously described in sections 3.3.3.1 and 3.3.3.2 of this chapter. This was a miscellaneous group that was called 'hormones other'. These eighteen probes were imported into GeneSpring software and following the procedure described previously (see section 3.2.2), a tree cluster heatmap, specific for these 18 probes, was created (Figure 3.8).

All the 15 probes belonging to the two clusters highlighted in the tree cluster heatmap of Figure 3.8 were reported on Tables 3.17, 3.18 and 3.19. The remaining three probes, those outside the two clusters highlighted in the tree cluster heatmap of Figure 3.8 were listed in Table A.3.5 (see appendix).

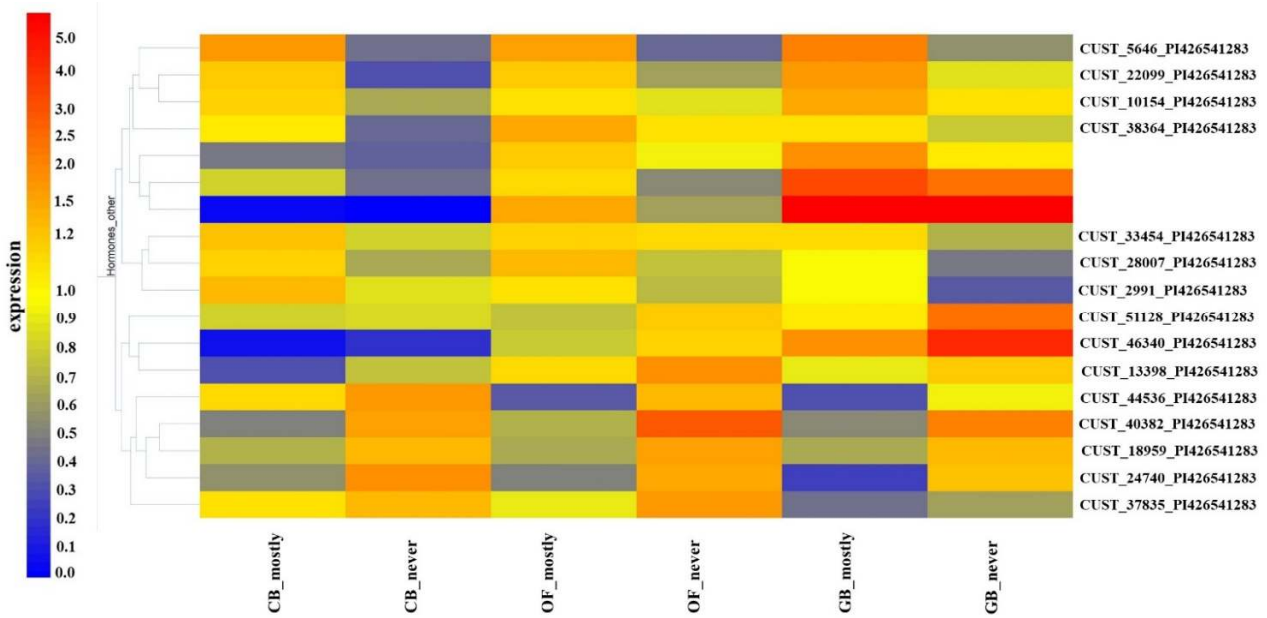


Figure 3.8: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to 'hormones other'.

Cluster tree heatmaps for the eighteen microarray probes differently expressed for what concerned the stage*phenotype interaction with differences being significative at 95% confidence levels. Three stages tested (i.e. closed bud (**CB**), open flower (**OF**) and green berry (**GB**)) and two phenotypes (i.e. mostly and never 'crumbly'). The eighteen probes matched *Rubus idaeus*' genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to any other hormone term, here for simplicity called 'hormones other', rather than the two previously presented in this chapter (i.e. 'response to hormone' and 'hormone biosynthesis'). Two main clusters of probes with evident difference in the expression level, as highlighted by the colour shades of the heatmap, were selected and the corresponding probes were reported on the right side of the heatmap. The four probes, within the first highlighted cluster, were upregulated in the mostly 'crumbly' phenotype compared to the never one. The last eleven probes of the heatmap, those within the second cluster, were in part again upregulated (i.e. the first three) in the mostly 'crumbly' phenotype and the remaining eight downregulated always in the mostly 'crumbly'. High expression levels are indicated in red colour while low expression levels in blue colour as per scale bar presented on the left side of the heatmap.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	linkage group
1st cluster	CUST_5646_PI426541283	[c]regulation of SA metabolic process	GO:0010337	AT1G28380.1	
	CUST_22099_PI426541283	steroid binding	GO:0005496	AT5G52240.1	[a]LG1
	CUST_10154_PI426541283	cytokinin metabolic process	GO:0009690	AT3G63440.1	[b]LG3_new
	CUST_38364_PI426541283	[d]SA mediated signalling pathway	GO:0009863	AT1G53130.1	
		[e]JA mediated signalling pathway)	GO:0009867		

[a]LG1 – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 1
[b]LG3_new – mapped inside the new 'crumbly' QTL identified here in this work on linkage group 3
[c]regulation of salicylic acid (SA) metabolic process
[d]salicylic acid (SA) mediated signalling pathway
[e]jasmonic acid (JA) mediated signalling pathway

Table 3.17: 'Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'hormone other'.

List of the four microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to the miscellaneous group 'hormone other'. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	linkage group
2nd cluster	CUST_33454_PI426541283	auxin activated signalling pathway	GO:0009734	AT1G05180.1	[a]LG3_old
		auxin homeostasis	GO:0010252		
	CUST_28007_PI426541283	[b]MeIAA esterase activity	GO:0080030	AT3G29770.1	
		[c]MeJA esterase activity	GO:0080032		
		[d]MeSA esterase activity	GO:0080031		
		salicylic acid metabolic process	GO:0009696		
		jasmonic acid metabolic process	GO:0009694		
	CUST_2991_PI426541283	[e]reg. of JA med. sig. pathway	GO:2000022	AT3G17860.3	
		JA mediated signalling pathway	GO:0009867		
	CUST_51128_PI426541283	sterol transporter activity	GO:0015248	AT3G09300.1	
		sterol binding	GO:0032934		
	CUST_46340_PI426541283	hormone activity	GO:0005179	AT2G16385.1	[a]LG3_old

[a]LG3_old – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3
[b]methyl indole-3-acetate esterase activity
[c]methyl salicylate esterase activity
[d]methyl jasmonate esterase activity
[e]regulation of jasmonic acid mediated signalling pathway

Table 3.18: : 'Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'hormone other'.

List of the first five, out of the eleven, microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to the miscellaneous group 'hormone other'. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	linkage group
2nd cluster	CUST_13398_PI426541283	abscisic acid (ABA) metabolic process	GO:0009687	AT2G04240.1	[a]LG3_old
	CUST_44536_PI426541283	ABA activated signalling pathway	GO:0009738	AT2G03440.1	
	CUST_40382_PI426541283	ABA activated signalling pathway	GO:0009738	AT3G51550.1	
		[b]brassinosteroids med. sig. pathway	GO:0009742		
		[c]neg. reg. of ABA-activated sig. path.	GO:0009788		
		[d]ethylene-activated sig. pathway	GO:0009873		
	CUST_18959_PI426541283	[e]gibberellic acid med. sig. pathway	GO:0009740	AT4G24210.1	
		[f]pos. reg. of gibb. acid med. sig. path.	GO:0009939		
	CUST_24740_PI426541283	sterol binding	GO:0032934	AT4G22540.1	
		signal transduction	GO:0007165		
		sterol transporter activity	GO:0015248		
	CUST_37835_PI426541283	Auxin-activated signalling pathway	GO:0009734	AT1G30330.2	[a]LG3_old

[a]LG3_old – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3
[b]brassinosteroids mediated signalling pathway
[c]negative regulation of abscisic acid (ABA) activated signalling pathway
[d]ethylene-activated signalling pathway
[e]gibberellic acid mediated signalling pathway
[f]positive regulation of gibberellic acid mediated signalling pathway

Table 3.19: 'Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'hormone other'.

List of the remaining six, out of the eleven, microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to the miscellaneous group 'hormone other'. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0010337 (regulation of SA metabolic process)	AT1G28380.1	mostly	0.263	0.221	0.358	0.0778	CUST_5446_PI426541283
		never	-0.35	-0.394	-0.25		
GO:005496 (steroid binding)	AT5G52240.1	mostly	0.082	0.08[d]	0.25[d]	0.1526	CUST_22099_PI426541283[b]
		never	-0.497	-0.187[d]	-0.044[d]		
GO:0009690 (cytokinin metabolic process)	AT3G63440.1	mostly	0.055	0.03[d]	0.205	0.0485	CUST_10154_PI426541283[c]
		never	-0.185	-0.044[d]	0.027		
[e]GO:0009863 (SA med. sign. path.) [f]GO:0009867 (JA med. sign. path.)	AT1G53130.1	mostly	0.017	0.186	0.024	0.0556	CUST_38364_PI426541283
		never	-0.39	0.024	-0.107		

[a]2 degrees of freedom and 4 replicates
[b]LG1 – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 1
[c]LG3_new – mapped inside the new 'crumbly' QTL identified here in this work on linkage group 3
[d]difference not statistically significant
[e]salicylic acid (SA) mediated signalling pathway
[f]jasmonic acid (JA) mediated signalling pathway

Table 3.20: ANOVA table of means for all the probes within the first cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'hormone other.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction for the four probes within the first cluster highlighted in the 'hormone other' genes tree cluster heatmap (Figure 3.8). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to the miscellaneous group 'hormone other', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009734 [b](auxin act. sign. path.) GO:0010251 (auxin homeostasis)	AT1G05180.1	mostly	0.097	0.07[c]	0.042	0.0441	CUST_33454_PI426541283[d]
		never	-0.077	0.037[c]	-0.152		
GO:0080030 [e](MeIAA esterase ctivity) GO:0080032 [f](MeJA esterase activity)	AT3G29770.1	mostly	0.059	0.135	-0.001	0.0907	CUST_28007_PI426541283[d]
		never	-0.176	-0.123	-0.33		
GO:2000022 [h](regulation of JA med.) GO:0009867 [i](JA mediated sign. path.)	AT3G17860.3	mostly	0.128	0.027	-0.006	0.0627	CUST_2991_PI426541283
		never	-0.046	-0.141	-0.441		
GO:0015248 (sterol transporter activity) GO:0032934 (sterol binding)	AT3G09300.1	mostly	-0.082[c]	-0.115	0.012	0.0483	CUST_51128_PI426541283
		never	-0.068[c]	0.081	0.411		
GO:0005179 (hormone activity)	AT2G16385.1	mostly	-1.32	-0.092[c]	0.288	0.0897	CUST_46340_PI426541283
		never	-0.723	0.071[c]	0.657		

[a]2 degrees of freedom and 4 replicates - [b]auxin activated signalling pathway - [c]difference statistically not significant
[d]LG3 – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3 - [e]methyl indole-3-acetate esterase activity
[f]methyl salicylate esterase activity - [g]methyl jasmonate esterase activity - [h]regulation of jasmonic acid mediated signalling pathway - [i]jasmonic acid mediated signalling pathway

Table 3.21: ANOVA table of means for all the probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'hormone other'.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction for the first five probes within the second cluster highlighted in the 'hormone other' genes tree cluster heatmap (Figure 3.8). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to the miscellaneous group 'hormone other', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009687 (ABA metabolic process)	AT2G04240.1	mostly	-0.503	0.048	-0.037[b]	0.07005	CUST_13398_PI426541283[c]
		never	-0.11	0.272	0.091[b]		
GO:0009738 [d](ABA activ. sign. path.)	AT2G03440.1	mostly	0.043	-0.462	-0.51	0.0693	CUST_44536_PI426541283
		never	0.239	0.146	-0.018		
GO:0009738 [d](ABA activ. sign. path.) GO:0009742 [e](brassinost. med. sig. path) GO:0009788 [f](neg. reg. of ABA act.) GO:0009873 [g](ethylene act. sig. path.)	AT3G51550.1	mostly	-0.295	-0.158	-0.262	0.1306	CUST_40382_PI426541283
		never	0.21	0.481	0.355		

[a]2 degrees of freedom and 4 replicates
[b]difference statistically not significant
[c]LG3 – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3
[d]ABA activated signalling pathway
[e]brassinosteroids mediated signalling pathway
[f]negative regulation of ABA activated signalling pathway
[g]ethylene activated signalling pathway

Table 3.22: ANOVA table of means for all the probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'hormone other'.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction for other three probes within the second cluster highlighted in the 'hormone other' genes tree cluster heatmap (Figure 3.8). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to the miscellaneous group 'hormone other', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

Gene Ontology (GO) term	A. thaliana gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009740 [d](gibb. acid med. sig. path.) GO:0009939 [e](positive reg. of gibb.)	AT4G24210.1	mostly	-0.148	-0.171	-0.185	0.0777	CUST_18959_PI426541283
		never	0.148	0.224	0.128		
GO:0015248 (sterol transporter activity) GO:0032934 (sterol binding) GO:0007165 (signal transduction)	AT4G22540.1	mostly	-0.244	-0.286	-0.573	0.0701	CUST_24740_PI426541283
		never	0.272	0.189	0.098		
GO:0009687 (auxin activated signalling pathway)	AT1G30330.2	mostly	0.021[b]	-0.037	-0.359	0.0646	CUST_37835_PI426541283[c]
		never	0.145[b]	0.241	-0.201		

[a]2 degrees of freedom and 4 replicates
[b]difference statistically not significant
[c]LG3 – mapped inside the 'crumbly' QTL identified previously by Graham *et al.,* (2015) on linkage group 3
[d]gibberellic acid mediated signalling pathway
[e]positive regulation of gibberellic acid mediated signalling pathway

Table 3.23: ANOVA table of means for all the probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'hormone other'.

ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction for the last three probes within the first cluster highlighted in the 'hormone other' genes tree cluster heatmap (Figure 3.8). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to the miscellaneous group 'hormone other', were reported. Moreover, if a gene was mapped inside one of the three 'crumbly' QTLs (see chapter 4 for details), it was reported in the table as was the linkage group in which the QTL was located. In the first column of the table was indicated the cluster of the heatmap to which the probes belong.

In Tables 3.20 to 3.23 the *Arabidopsis thaliana* gene IDs for the 15 ortholog genes, of those *Rubus idaeus*, matched by the 'crumbly' microarray experiment probes were reported together with the predicted means with their standard error from the analysis of variance (ANOVA). These fifteen probes were included in the two highlighted clusters of the heatmap of Figure 3.8; these two clusters contained potential interesting genes for the 'crumbly' fruit whose gene ontology annotations were related to a miscellaneous group of hormones related annotation (Tables 3.17, 3.18 and 3.19) here called 'hormones other' for simplicity.

The four *Arabidopsis thaliana* gene orthologs of the *Rubus idaeus*' ones matched by the microarray probes, had five different gene ontology annotations. Two of these ontology terms refer to signalling pathways mediated by salicylic and jasmonic acid. The activation of a molecular signal pathway mediated by these two hormones that conveys stimuli to a recipient becomes particularly interesting, according to the hormonal crosstalk hypothetical fruit growth regulating process that forms part of the rationale behind this work. The idea is of a hormonal crosstalk between the receptacle, the leading hub, and fertilised ovaries to help synchronise the growth of all the different parts of the raspberry fruit. The receptacle by sending molecular messages, through the activation of hormonal signalling pathways, helps to regulate/coordinate fruit growth. The other ontology annotation refers to salicylic acid and cytokinin metabolism and to steroid binding; the genes with ontology related to cytokinin and steroid were mapped inside two of the three 'crumbly' QTLs (see Tables 3.17, 3.18 and 3.19) giving further evidence for to their potential involvement in generating 'crumbly' fruit. The first cluster would, from the gene ontology annotations perspective, represent the effect of four different hormones (i.e. salicylic acid, jasmonic acid, cytokinin and steroid) in potentially causing 'crumbly' fruit. All four genes were upregulated in the mostly 'crumbly' phenotype compared to never 'crumbly' indicating that the overexpression of these gene might contribute to the formation of 'crumbly' like fruit in red raspberry.

The function of the four genes identified in the first cluster of the heatmap in Figure 3.8 were:

- AT1G28380.1 - encodes the NECROTIC SPOTTED LESIONS 1 (NSL1), a predicted protein involved in negatively regulating salicylic acid-related defense responses and cell death programs (source TAIR - https://www.arabidopsis.org/).

- AT5G52240.1 – encodes a Membrane Steroid Binding Protein 1, (MSBP1) expressed in flower carpel and pollen. The protein MSBP1 functions as a negative regulator of cell elongation in Arabidopsis; overexpression of *MSBP1* affects the expression of specific genes involved in cell elongation and sterol metabolism (Yang et al., 2005).
- AT3G63440.1 – as described in the previous section encodes Cytokinin oxidase 6 (CKX6 previously called CKX7) that catalyses the irreversible inactivation of CKs, transforming them to adenine and a corresponding side chain-derived aldehyde. In fruit CKX7 degrades CKs inhibiting cell elongation (Di Marzo et al. 2020).
- AT1G53130.1 – encodes GRIM REAPER (GRI), involved in the regulation of cell death induced by extracellular ROS (reactive oxygen species). The gene is expressed in flower and vascular leaf (source TAIR - https://www.arabidopsis.org/).

The analysis of the functions, related to the four genes belonging to the first cluster of differentially expressed microarray probes of Figure 3.8, showed interesting results related to 'crumbly' fruit, particularly for two of the genes considered (i.e. AT5G52240.1 and AT3G63440.1). The first was involved in molecular processes that negatively affect cell elongation which is an important step that takes place during fruit growth and development. The microarray probe matching the gene showed significant differences in the expression level between the two phenotypes (i.e. mostly and never 'crumbly') only at closed bud stage (Table 3.20). Impairment in this specific biological process might then affect flower growth and therefore impact on fruit development. AT5G52240.1 encodes the protein MSBP1 that functions as a negative regulator of cell elongation; overexpression of *MSBP1* affects the expression of specific genes involved in cell elongation (Yang at al. 2005). The gene AT3G63440.1 positively affect fruit elongation because it encodes a cytokinin dehydrogenase that degrades cytokinin, an important inhibitor of cell elongation during fruit development (Di Marzo et al. 2020). The differences in the expression levels between the two phenotypes (i.e. mostly and never 'crumbly') were significant only at the closed bud stage for AT5G52240.1 and at both closed bud and green stage for AT3G63440.1 (see Table 3.20). Both these genes were upregulated in the mostly 'crumbly' plants and both genes were mapped inside two of the 'crumbly' QTLs, the one on linkage group 1 previously identified by Graham et al. (2015) and the new QTL on linkage group 3 identified during this work (see chapter four for more details). Such findings support the roles of these genes in terms of their contribution,

direct or indirect, to the formation, growth and development of misshapen 'crumbly' like raspberry fruits

The second cluster of genes, identified in Figure 3.8, contained eleven genes with twenty-four different gene ontology annotations. Eleven of these ontology terms refer to hormonal signalling pathways that may indicate the importance of such signal transduction in fruit development. Such molecular signals would trigger changes in the activity or state of a cell that are important for regulating/controlling downstream processes (e.g. regulation of transcription or regulation of a metabolic processes). The presence of so many gene ontology annotations related to hormonal signalling pathway, gives further support to the hypothesis, behind this work, about a hormonal regulating crosstalk between two different and separate parts of the flower, first, and then of the fruit (i.e. receptacle and ovaries) that permits the synchronised growth of all the many fertilised ovaries that compose the raspberry fruit. Two annotations refer to sterols transport, which means the directed movement of sterols into, out of or within a cell, or between cells; sterols are precursors of steroids (i.e. Brassinosteroids being the most well characterised). The same consideration, related to signalling pathways, also apply to the ontology term (sterol transport). All the remaining annotations terms concern hormones metabolism. This second cluster, highlighted on the tree cluster heatmap (Figure 3.8), might represent hypothetical hormonal interplay that, according to the hypothesis behind this work, allows the synchronised growth of all the fertilised ovaries and permits the correct growth and development of the raspberry fruit.

The function of the eleven genes identified in the second cluster of the heatmap in Figure 3.8 were:

- AT1G05180.1 – encodes AUXIN RESISTANT 1 (AXR1) a subunit of the RUB1 activating enzyme that regulates the protein degradation activity of Skp1-Cullin-Fbox complexes, primarily, but not exclusively, affecting auxin responses. AXR1 gene is required for auxin action in most, if not all, tissues of the plant and plays an important role in plant development (Lincoln et al., 1990).
- AT3G29770.1 – encodes a Methyl Esterase 11 (MES11), involved in jasmonic and salicylic acid metabolic process and with methyl indole-3-acetate, methyl jasmonate and methyl salicylate esterase activity. The enzyme is expressed in flowers (source TAIR - https://www.arabidopsis.org/)

- AT3G17860.3 – encodes the Jasmonate-Zim-Domain protein 3 (JAZ3). The protein participates in the jasmonic acid signalling pathway. After its synthesis inside the peroxisome, the jasmonic acid is moved into the cytoplasm, where it is conjugated with isoleucine (Ile) to produce bioactive JA-Ile. In the JA signalling pathway, JA-Ile promotes the interaction between COI1, a JA receptor, and the JAZ protein. At this point, JAZ can be ubiquitinated and degraded by the 26S proteasome, leading to the release of MYC2, the major transcription factor of JA-mediated genes with consequent activation of the genes responsive to JA (Huang et al., 2017).
- AT3G09300.1 – encodes the OSBP (oxysterol binding protein)-related protein 3B (ORP3B) with sterol binding and sterol transporter activity. The protein is expressed in flower and pollen (source TAIR - https://www.arabidopsis.org/)
- AT2G16385.1 – encodes the Casparian Strip Integrity Factor 1, (CAF1) a peptide hormone known to be expressed in the root stele that specifically binds the endodermis-expressed leucine-rich repeat receptor kinase GASSHO1 (GSO1)/SCHENGEN3 and its homolog, GSO2. Together with CAF2 it is required for formation of the casparian band (source TAIR - https://www.arabidopsis.org/).
- AT2G04240.1 – encodes XERICO, a small protein with an N-terminal trans-membrane domain and a RING-H2 zinc finger motif located at the C-terminus. The transcription of the gene is induced by DELLA proteins and repressed by gibberellic acid; the protein is involved in ABA metabolism (source TAIR - https://www.arabidopsis.org/).
- AT3G51550.1 – encodes FERONIA (FER) a synergid-expressed, plasma-membrane localized receptor-like kinase that accumulates asymetrically in the synergid membrane. For successful fertilization, the reception of the pollen tube inside the embryo sac is crucial and FERONIA seems to be the determinant for the correct interaction between the pollen tube and the ovule (Haruta et al., 2018b). The FERONIA enzyme seems indispensable for the arrest of the pollen tube growth and its release of the sperm cells and to guarantee the ovule fertilization.
- AT4G24210.1 – encodes SLEEPY1 (SLY1), a F-box protein that is involved in GA signalling. SLEEPY1 is involved in DELLA proteins which are the main negative regulator of gibberellin signalling pathway. SLEEPY1 by modulating DELLA impact on gibberellins response genes with potential effect, above all, on plant flowering and fertility (McGinnis et al., 2003).

- AT4G22540.1 –encodes the OSBP (oxysterol binding protein)-related protein 3B (ORP2A) with sterol binding and sterol transporter activity. The protein is expressed in flower and pollen (source TAIR - https://www.arabidopsis.org/)
- AT1G30330.2 – encodes ARF6, that as showed in Liu et al., (2018) is responsible for the correct maturation of stamens and for plant fertility.

The analysis of the molecular functions fulfilled by the eleven genes located inside the second cluster of probes for the genes with ontology annotation related to 'hormones other' (Figure 3.8), again highlighted the importance of flower differentiation, pollen formation and early stage of ovule fertilization as key processes for 'crumbly' fruit development. The analysis of the genetic expression of plants showing 'crumbly' phenotype, once again showed how impairment in processes such as stamen maturation, pollen formation and interaction pollen tube ovules, as for the three genes (i.e. AT3G51550.1, AT4G24210.1 and AT1G30330.2) could determine a reduction in the fertilization rate with consequent formation of fruit with a lower number of drupelets that might therefore be misshapen. It can be concluded that, from a 'crumbly' fruit perspective, this second cluster would represent the impairment in the pollen formation and maturation as well as in the pollen tube/ovule interaction.

3.3.4 Heatmap of microarray probes matching genes with GO terms related to transport

The microarray analysis highlighted twenty probes, significantly differentially expressed between the mostly and the never 'crumbly' phenotypes, with specific gene ontology terms related to 'transport'. These twenty probes were imported into GeneSpring software and following the procedure described in section 3.2.2, a tree cluster heatmap, specific for these 20 probes, was created (Figure 3.9). Two main cluster of probes were highlighted respectively on the centre and bottom of the heatmap and the corresponding probes, within each cluster, were reported on the right side of the heatmap (see Figure 3.9).

The full list of the 9 probes belonging to the two clusters highlighted in the tree cluster heatmap of Figure 3.9 was reported on Table 3.24. The remaining 11 probes, those outside the two clusters, always of the tree cluster heatmap of Figure 3.9 were listed and reported in appendix (see Table A.3.6). For each probe the *Arabidopsis thaliana* ortholog gene to

the *Rubus idaeus*' one, matched by the microarray probe, was reported together with the gene ontology term, obviously related to 'transport', and its corresponding code.

In Table 3.24, the 12 *Arabidopsis thaliana* gene IDs ortholog of those *Rubus idaeus*' one matched by the 'crumbly' microarray experiment probes were reported. The table included only those probes contained in the two highlighted clusters of potential interesting genes for the 'crumbly' fruit that have gene ontology annotations related to 'transport'. For each probe, in Tables 3.25 and 3.26, the predicted means with their standard error from the analysis of variance (ANOVA), stage*phenotype interaction, to show whether for each stage, the difference in the expression level of each probe was statistically significative at 95% confidence levels were reported.

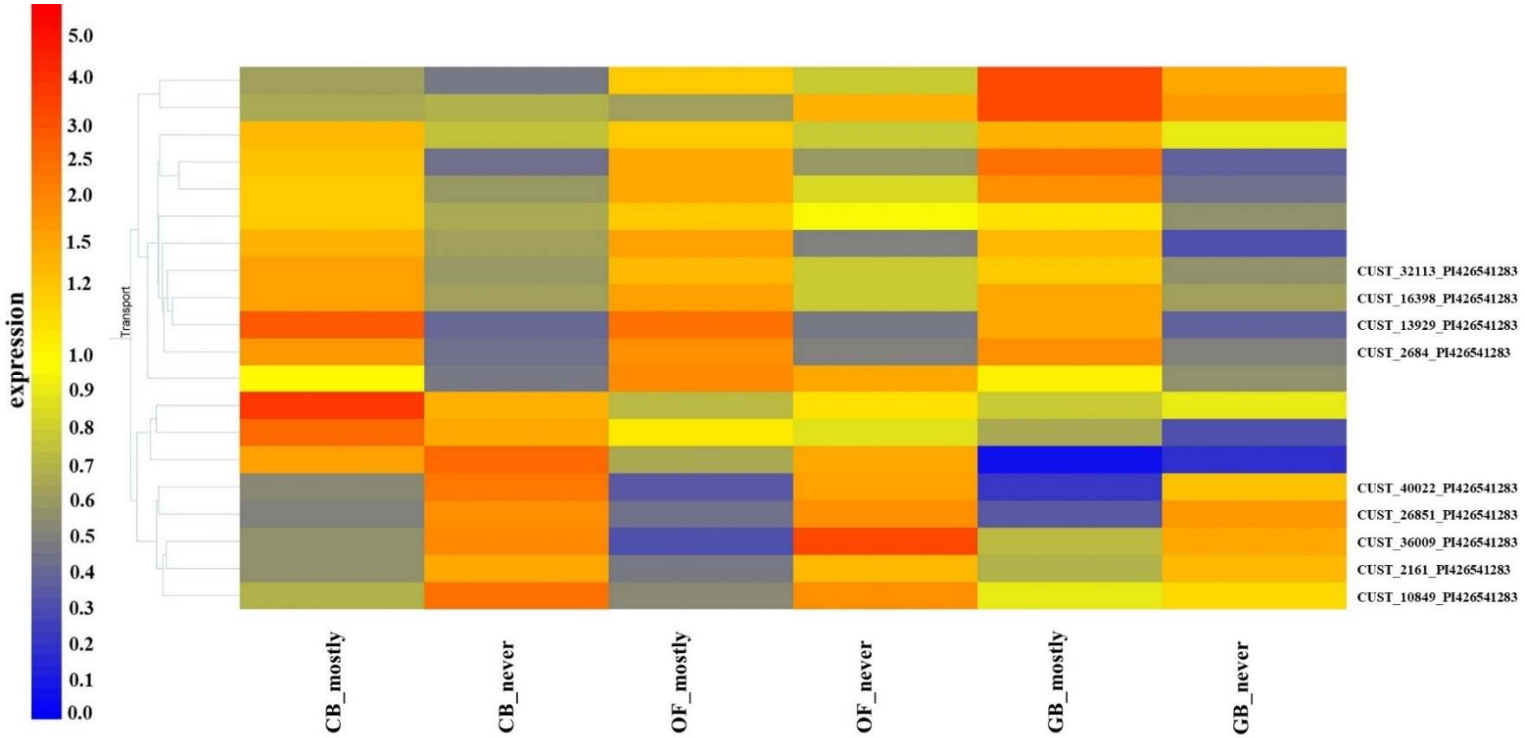


Figure 3.9: Tree cluster heatmap of the 'crumbly' microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology terms related to transport.

Cluster tree heatmaps for the twenty microarray probes differently expressed for what concerned the stage*phenotype interaction with differences being significative at 95% confidence levels. The twenty probes matched *Rubus idaeus*' genes whose *Arabidopsis thaliana* orthologs had gene ontology annotations related to 'transport' terms. Two main clusters of probes with evident difference in the expression level, as highlighted by the colour shades of the heatmap, were selected and the corresponding probes were reported on the right side of the heatmap. The four probes, within the first highlighted cluster, were upregulated in the mostly 'crumbly' phenotype compared to the never one. The last five probes of the heatmap, those within the second cluster, were downregulated in the mostly 'crumbly' phenotype. High expression levels are indicated in red colour while low expression levels in blue colour as per scale bar presented on the left side of the heatmap.

	Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID
1st cluster	CUST_33454_PI426541283	transmembrane transport	GO:0055085	AT3G56950.1
	CUST_16398_PI426541283	cation transmembrane transporter activity	GO:0008324	AT2G04620.1
	CUST_13929_PI426541283	monovalent cation:proton antiport activity	GO:0005451	AT1G79400.1
	CUST_13929_PI426541283	intracellular protein transport	GO:0006886	AT3G24390.1
2nd cluster	CUST_40022_PI426541283	transmembrane transport	GO:0055085	AT1G51610.1
		efflux transmembrane transporter activity	GO:0015562	
	CUST_26851_PI426541283	calcium ion transport	GO:0006816	AT5G53130.1
	CUST_36009_PI426541283	peptide transmembrane transporter activity	GO:1904680	AT1G69870.1
		nitrate transport	GO:0015706	
	CUST_2161_PI426541283	polysaccharide transport	GO:0015774	AT2G18840.1
	CUST_10849_PI426541283	amino acid transmembrane transport	GO:0003333	AT2G41190.1

Table 3.24: 'Crumbly' microarray probes matching A. thaliana genes with ontology terms related to 'transport'.

List of the nine microarray probes significantly differently expressed, with differences being significant at 95% confidence levels. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; related to 'transport'. In the first column of the table was indicated the cluster of the heatmap (Figure 3.9) to which the probes belong.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0055085 (transmembrane transport)	AT3G56950.1	mostly	0.231	0.128	0.088	0.0692	CUST_32113_PI426541283
		never	-0.215	-0.105	-0.251		
GO:0008324 [b](cation transmem. transp. act.)	AT2G04620.1	mostly	0.209	0.221	0.181	0.0493	CUST_16398_PI426541283
		never	-0.196	-0.1	-0.192		
GO:0005451 [c](monovalent cation:proton ant.)	AT1G79400.1	mostly	0.493	0.4	0.203	0.1166	CUST_13929_PI426541283
		never	-0.396	-0.33	-0.414		
GO:0006886 (intracellular protein transport)	AT5G24390.1	mostly	0.244	0.27	0.278	0.0611	CUST_2684_PI426541283
		never	-0.341	-0.308	-0.305		

[a]2 degrees of freedom and 4 replicates
[b]cation transmembrane transporter activity
[c]monovalent catio:proton antiporter activity

Table 3.25: ANOVA table of means for all the probes within the first cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'transport'. ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction for the last three probes within the first cluster highlighted in the 'transport' genes tree cluster heatmap (Figure 3.9). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'transport', were reported.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0055085 [c](transmem.trans.) GO:0015562 [d](efflux transm. tr.)	AT1G51610.1	mostly	-0.259	-0.469	-0.659	0.1242	CUST_40022_PI426541283
		never	0.368	0.236	0.099		
GO:0006816 (calcium ion transport)	AT5G53130.1	mostly	-0.29	-0.353	-0.467	0.1225	CUST_26851_PI426541283
		never	0.272	0.284	0.251		
GO:1904680 [e](peptide transm.) GO:0015706 (nitrate transport)	AT1G69870.1	mostly	-0.235	-0.501	-0.14	0.0827	CUST_36009_PI426541283
		never	0.301	0.554	0.206		
GO:0015774 (polysaccharide transport)	AT2G18840.1	mostly	-0.231	-0.33	-0.155	0.0632	CUST_2161_PI426541283
		never	0.19	0.124	0.123		
GO:0003333 [f](amino acid transm. transport)	AT2G41190.1	mostly	-0.147	-0.277	-0.036[b]	0.0641	CUST_10849_PI426541283
		never	0.404	0.276	0.051[b]		

[a]2 degrees of freedom and 4 replicates
[b]difference not statistically significant
[c]transmembrane transport
[d]efflux transmembrane transporter activity
[e]peptide transmembrane transporter activity
[f]amino acid transmembrane transport

Table 3.26: ANOVA table of means for all the probes within the second cluster highlighted in the tree cluster heatmap for the genes with gene ontology related to 'transport'. ANOVA of the normalised (i.e. logarithm of base 2) expression levels for the stage*phenotype interaction for the last three probes within the second cluster highlighted in the 'transport' genes tree cluster heatmap (Figure 3.9). For each probe were reported the predicted means from the analysis of variance and the standard error to show how the differences in the expression levels, if any, were significative at 95% of confidence levels for the two different phenotypes (i.e. mostly and never 'crumbly') analysed, in all the three stages tested (i.e. closed bud, open flower and green berry). For each probe, the matched Arabidopsis thaliana gene ID and its gene ontology annotations, related to 'transport', were reported.

For the first cluster of probes in Figure 3.9, the four *Arabidopsis thaliana* gene orthologs of the *Rubus idaeus*' ones matched to the microarray probes, had four different gene ontology annotations. In order to help select the best annotations; the aim was to identify potential genes whose function could be related to the transport/movement of phytohormones as in the proposed putative hormonal crosstalk between receptacle and fertilised ovaries hypothesized during this work (see chapter 1 for more details).

The gene ontology resource database (http://geneontology.org/) was consulted for the correct definitions to help select the best ontology terms that could explain the existence of a putative hormonal interplay responsible for the control/regulation of the fruit growth and development. Focusing again on the ontology terms related to the four genes within the first cluster for the heatmap in Figure 3.9, the first annotation is 'transmembrane transport' whose definition is: the process in which a solute is transported across a lipid bilayer, from one side of a membrane to the other. The second and third ontology terms, 'cation transmembrane transporter activity' and 'monovalent cation/proton antiporter activity', indicates respectively the ability to transfer a cation from one side of a membrane to the other. While the second indicates the intervention of a carrier protein, an antiport, which is essential to import a monovalent cation required to export another monovalent cation, according to the schematised reaction: monovalent cation(out) + H^+(in) = monovalent cation(in) + H^+(out). The fourth gene ontology annotation was 'intracellular protein transport' which refers to the directed movement of proteins inside a cell, including the movement of proteins between specific compartments or structures such as organelles. All these ontology terms were very generic and so no conclusion could be made about their putative involvement in processes related to the growth and formation of misshapen 'crumbly' like fruit. All the probes within this first cluster, and consequently their matched genes, were upregulated in the mostly 'crumbly' phenotype.

The function of the four genes identified in the first cluster of the heatmap in Figure 3.9 were:

- AT3G56950.1 – encodes a protein belonging to a subfamily of aquaporins, the Small and basic Intrinsic Protein 2 (SIP2;1) which facilitate the passive transport of small molecules across membranes. The protein is localised in the endoplasmic reticulum (ER) and mutants (sip2;1). Although having normal vegetative growth, these mutants

showed reduced pollen germination rate (about 60% compared to wild type plants), reduced elongation length of the pollen tube (up to 80% compared to wild type) and consequently even the fertilization rate was affected (Sato and Maeshima, 2019). The pollen itself, presented normal shape, was viable, swelled after pollination, and adhered correctly to stigma. The ER synthesizes proteins and other molecules for the plasma membrane (PM). It generates small intracellular membrane vesicles that contain the synthesised PM components and the cell wall precursors. These membrane vesicles deliver the components to the PM from the ER through the Golgi apparatus. In wild type plants, SIP2;1 is evenly distributed in the ER of growing pollen tube (Sato and Maeshima, 2019). It might be speculated that in the mutants (sip2;1) a latent uneven distribution of the SIP2;1 protein on the ER might negatively affect the delivering of PM components and cell wall precursor and this, particularly during pollen tube protrusion, when there is the highest requirement for those components would compromise pollen tube elongation and consequently ovule fertilisation.

- AT2G04620.1 – encodes (MTP12), a Zn transporter that is localized on the Golgi apparatus and forms a functional complex with AtMTP5t1 to transport Zn into the Golgi (source TAIR - https://www.arabidopsis.org/).
- AT1G79400.1 – encodes Cation/H^+ Exchanger 2 (CHX2) with monovalent cation/proton antiporter activity and a sodium/proton antiporter activity. It is expressed in mature pollen and germinated pollen stages (source TAIR - https://www.arabidopsis.org/).
- AT5G24390.1 – encodes the Ypt/Rab-GAP domain of gyp1p superfamily protein, a intracellular transport protein (source TAIR - https://www.arabidopsis.org/).

The selection of genes with gene ontology annotations related to 'transport' was performed in order to try to identify any putative mechanism that could explain or at least have reference to the fruit development in which an hormonal interplay between different parts of the fruit/flower was proposed. Such hormonal crosstalk would occur through the exchange of molecular messages, for instance through a transportation system, between receptacle and the fertilised ovaries. An important finding, from a 'crumbly' fruit perspective, was for the gene AT3G56950.1. It encodes SIP2;1 an aquaporin located on the membrane of the endoplasmic reticulum (ER). Mutants (sip2;1) showed defects in pollen germination, pollen tube elongation and subsequent ovule fertilization. The gene was differently expressed between the two phenotypes (i.e. mostly and never 'crumbly')

in all the three tested stage (i.e. closed bud, open flower and green berry). It might be the case that the difference in the expression level in mostly 'crumbly' might produce an effect similar to that of the (sip2;1) mutants (Sato and Maeshima, 2019) where plants producing normal shape pollen able to adhere to the stigma but unable to fertilise were identified. This would have a consequence for the reduction in the number of fertilised ovaries that would give rise to misshapen fruits with lower number of drupelets such as those found in 'crumbly' plants.

The second cluster of probes, highlighted in the heatmap of Figure 3.9, had seven different gene ontology annotation related to 'transport'. The full list of ontology terms was reported on Table 3.24. The cluster, in terms of the gene ontology annotation, represents the transport across the membrane, within a cell and between cells, of both small ions (i.e. calcium and nitrate) and of large molecules (i.e. amino acids, peptides and polysaccharides) by means of transporters or membranes pores. No direct connection was identified with putative transport mechanisms involving plant hormones except for the genetic annotation 'efflux transmembrane transporter activity' which refers to a general movement of solutes outside a cell. It cannot be excluded a priori that such system would be used by plants to convey plant hormones from the site of synthesis/activation to that of action. Regardless the fact that all these five genes were differently expressed in the mostly 'crumbly' phenotype compared to the never one would indicate a potential role on 'crumbly' fruit development as all the genes were downregulated in the mostly 'crumbly' phenotype.

The functions of the five genes identified in the second cluster of the heatmap in Figure 3.9 were:

- AT1G51610.1 – encodes a cation efflux family protein involved in transmembrane transport and located in the chloroplast. The gene is widely in various tissues including carpel, stamen, flower and pollen (source TAIR - https://www.arabidopsis.org/).
- AT5G53130.1 – encodes Cyclic Nucleotide Gated Channel 1 (CNGC1), a member of the Cyclic nucleotide gated channel family that is involved in calcium and potassium ions transport. The gene is widely expressed including in carpel, stamen, flower and pollen (source TAIR - https://www.arabidopsis.org/).

- AT1G69870.1 – encodes the Nitrate Transporter 1.7 (NRT1.7), the gene is expressed in phloem. The protein is responsible for source-to-sink remobilization of nitrate and its mRNA is cell-to-cell mobile (source TAIR - https://www.arabidopsis.org/).
- AT2G18840.1 – encodes the YPT/RAB GTPase Interacting Protein 4a (YIP4a) which form a TGN-localized complex with ECHIDNA (ECH). This complex is required for the secretion of cell wall polysaccharides. The secretion of cell wall polysaccharides through the *trans*-Golgi network (TGN) is required for plant cell elongation (Gendre et al., 2013). The gene is expressed in pollen tube cells and the pollen tube is the fastest growing tissue in plants requiring cell wall components to cope with the high elongation rate during the protrusion of the pollen tube inside the flower style. The Golgi apparatus, via the *trans*-Golgi network (TGN), supplies the cell wall of the growing (elongating) cells with the extra polysaccharides they need. One of the components of the TGN is (YIP4a), encoded by AT2G18840.1, the gene was downregulated in the mostly 'crumbly' plants across all the three development stages (i.e. closed bud open flower and green berry) with difference being statistically significative (see Table 3.26). It was reasonable to speculate that the under-expression of these genes might affect pollen tube elongation compromising flower fertilisation with consequent formation of fruit with lower number of drupelets that might be misshapen like the 'crumbly' one.
- AT2G41190.1 – encodes a transmembrane amino acid transporter family protein, located in plasma membrane, nucleus and vacuole. The protein is involved in amino acid transmembrane transport and it is expressed, among many, in pollen and stamen (source TAIR - https://www.arabidopsis.org/).

The analysis of the functions of the five genes, within the second cluster of differentially expressed microarray probes from Figure 3.9, cannot be directly linked with either 'crumbly' fruit in general or to a potential mechanism linked to the transport of hormones thus currently their role remains unclear. The importance of a mechanism revealing the existence of a hormones transport supports the proposed idea of a hormonal crosstalk between receptacle and fertilised ovaries whose aim would be the control/regulation and synchronisation of the growth of the many ovaries that together compose a raspberry fruit. Despite that, one gene (AT2G18840.1) proved to be very interesting from a 'crumbly' fruit perspective, in fact it encodes a transporter located in the Golgi apparatus that is important for conveying cell walls components during cell elongation, the cell having to increase its size requires new

cell wall components to support the expansion. This is particularly important in high expanding tissues such as those of the pollen tube. The gene AT2G18840.1 is expressed in both pollen and pollen tube. In the mostly 'crumbly' phenotype, the gene was downregulated suggesting that a reduction in the expression level of this gene might cause impaired fertilization since the pollen tube cannot elongate enough inside the style with consequent missed fertilization. A fruit with lower number of fertilised ovaries is more probable to be misshapen like in the 'crumbly' fruit.

3.3.5 Summary

The analysis of the tree cluster heatmaps, generated using the microarray probes differentially expressed between the two phenotypes (i.e. mostly and never 'crumbly'), was used to identify genes controlling molecular functions that would have a potential role in processes that could cause the formation of misshapen 'crumbly' like fruit. Obviously, the analysis of all the differently expressed microarray probes (827 in total) would have been impracticable and so the study was focused on four different groups each representing a different gene ontology class (i.e. hormones, flower development, pollen and transport). The choice to focus on specific gene ontology annotations such as flower development and pollen was dictated by literature sources suggesting that 'crumbly' fruit might be triggered by anomalies affecting embryo sac development and pollen fertility (Jennings 1988). It seemed then reasonable to look at genes with putative functions related to pollen and flower development. Genes with ontology terms related to hormones were chosen because these molecules are by far the main plant growth regulators therefore their involvement in the processes leading to fruit growth and development cannot be excluded. The fourth class of ontology annotations (i.e. transport) was contemplated because a putative system of transportation, to move hormones from one side to another or to convey molecular signals induced by hormones, is the basis of the hormonal interplay, between receptacle and fertilised ovaries, here proposed as the mechanism developed by plants to synchronise the growth of all the fertilised ovaries that compose a raspberry fruit.

From the analysis of the molecular functions, for those genes selected through the tree cluster heatmaps (see section 3.3), 26 genes were selected. The products of these genes are directly or indirectly involved in processes that could affect the normal, set, growth

and development of the flowers first and then of the fruits therefore potentially causing 'crumbly' fruit. Twenty of these genes were related to pollen and of half of these 'pollen' genes were required in those processes of synthesis or conveying of cell wall components. Anomalies in these processes can compromise the pollen tube growth because the cells cannot cope with the high demand of components required by the cell wall to sustain the expansion of the cell. Consequence of this disruption can result in the inability of the pollen tube to grow enough to reach the ovules. In fact, anomalies in cell elongation of the pollen tube during its protrusion inside the style represented by far, according to the results from the microarray experiment, the most abundant altered phenomena to potentially cause 'crumbly' fruit. To give more credit to this findings, two of the gene involved in these processes were mapped inside the 'crumbly' QTLs on linkage group 1 and 3. The remaining genes related to pollen are involved in, pollen formation, recognition of pollen on the stigma, germination and fertilization. Essentially the selected 'pollen' genes participate to all the principle stages regarding the pollen life cycle, from its formation until its eruption into the ovules. All these genes were differentially expressed in the 'mostly' crumbly phenotype indicating that disruption in those molecular processes could play a part in causing the formation of 'crumbly' like misshapen fruits in red raspberry.

Four genes, out of the 26, selected for their specific molecular function, are implicated in processes related to flower development and in particular to that of the stamens; again these genes were differently expressed between the two phenotypes and so anomalies in their functions might cause the formation of 'crumbly' fruit. Two of these genes mapped inside the 'crumbly' QTLs on linkage group 3 strengthening their position as potentially important players in the processes causing 'crumbly' fruit. The selected genes, as stated before, are implicated in process that affect primarily the development of the stamens. These last ones are the male flowers and the pollen-producing reproductive organ. Such aspect would indicate once more the importance of the disrupted process linked to pollen as key players in the 'crumbly' fruit phenomenon.

The remaining genes were related to hormones. Two of them are implicated in hormonal signalling and specifically in abscisic acid (ABA) and jasmonic acid (JA). This finding was very important because it gave some support to the hypothesis of a hormonal crosstalk responsible for the correct growth of the fruit. Impairment on this putative hormonal interplay might cause 'crumbly' fruit. The putative role played by ABA, in potentially causing 'crumbly' fruit, was strengthen by the results of the metabolomic

analysis (see chapter 5 for more details) where significantly higher concentrations of ABA were measured in receptacle samples of artificially induced 'crumbly' fruit. The last gene encodes for an enzyme responsible for the degradation of the cytokinin that are the main cell elongation repressor during fruit growth. Once more, from a 'crumbly' fruit perspective, impairment in the function of this gene might alter normal fruit growth with consequent potential formation of misshapen fruit.

All these findings would suggest that the 'crumbly' fruit phenotype, from an anatomical/molecular perspective, might be caused by a series of phenomena/processes affecting, above all, important steps of male gametophytes life cycle and in particular the cell elongation in the pollen tube during its protrusion inside the style of the carpel.

3.4 Discussion

Of the 827 genes differentially expressed with respect to the stage*phenotype interaction, one hundred and six were selected for their GO term related to four specific annotations: flower development, hormones, pollen and transport. These 106 were analysed with tree cluster heatmaps which allowed a further selection of sixty-nine genes that may have important roles in the 'crumbly' fruit phenotype. Of these sixty-nine genes, thirteen were related to pollen and involved in regulating important functions such as, pollen development and formation, pollen tube growth and pollen recognition/interaction with the ovule. The seven genes selected with GO related to flower development, were involved in the general development of the flower and in that of the embryo sac, in the double fertilization and in the development of gametophytes. The forty genes selected with GO terms related to hormones were divided in three groups, the first with genes activated in response to abscisic acid (ABA), Auxin, salicylic acid (SA), jasmonic acid (JA), cyclopentenone (OPDA) and cytokinins (Cks). The second group containing genes involved in the biosynthesis of sterol (precursor of brassinosteroids), JA, cinnamic acid (both a growth regulator and a precursor of SA), indole-3-acetate (IAA) and Cks. While the last group had genes involved in the metabolism of: SA, Cks, IAA, ABA and JA as well as in the signalling pathways of: ethylene, ABA, brassinosteroids (Brs), JA and gibberellins (GAs).

The analysis of the gene functions for the sixty-nine genes selected with the tree cluster heatmaps showed that twenty-six were those more critical in terms of 'crumbly' fruit. These twenty-six genes were divided in three groups. The majority of them belonged to a group of genes whose products were implicated in different aspects of pollen life cycle, from its formation until it erupts inside the ovule to release the two sperm cells. Four other genes were involved in processes linked with flower development. Anomalies in the functions of these genes might be the cause of the development of misshapen 'crumbly' like fruit.

Gibberellins (GAs) are important phytohormones regulating plant growth. They are involved in different development processes, with pollen maturation and flower induction among those (Daviere and Achard, 2013). Impairments of both GA biosynthesis and signalling can cause dwarfism, delayed or failed flowering and reduce fertility (McGinnis et al., 2003) (Kim et al., 2015). GA mediated development processes are negatively controlled, at the transcriptional level, by DELLA proteins which act as a repressor.

Sumoylation is a mechanism plants developed to modulate the repression activities of DELLA proteins by addition of ubiquitin, a conjugation reaction, in a process catalysed by the E3 SUMO (Small Ubiquitin-mediate MOdifier) ligase. In *Arabidopsis thaliana*, AtSIZ1 is the enzyme acting as E3 SUMO ligase, it targets different proteins and within those, SLEEPY1 a F-box protein, which is encoded by SLY1 gene (AT4G24210.1). Once sumoylated, SLEEPY1 becomes active and forms the $SCF^{SLY\text{-}SUMO}$ complex that interacts directly with DELLA proteins and ubiquinates them; ubiquitin conjugated DELLA proteins are then degraded by the 26S proteasome complex (Kim et al., 2015). In this study the SLY1 gene was shown to be downregulated in the mostly 'crumbly' phenotype and the differences in the expression levels were significant at the 95% confidence level in all the three stages (see Table 3.23 for further details). However, the highest difference was found in the open flower stage suggesting that in the mostly 'crumbly' phenotype a reduction in the GA response due to impaired modulation of DELLA proteins might affect the flowering as well as the pollen maturation, both processes that may play a crucial role in the formation and development of misshapen 'crumbly' like raspberry fruit.

The *Arabidopsis thaliana* gene AT3G51550.1 encodes a synergid-expressed, plasma-membrane localized receptor-like kinase that accumulates asymmetrically in the synergid membrane at the filiform apparatus and mediates male-female gametophyte interactions during pollen tube reception. Unlike animals where the product of meiosis are gametes, in plants, meiotic processes results in the formation of gametophytes pollen (male) and embryo sac (female). For fertilization, the pollen, after germination on the stigma, grows inside the flower style; the result of this growth is the so-called pollen tube that takes the two sperm cells (male gametes) to the embryo sac. In order to guarantee efficient delivery, chemical signals are used to guide the pollen tube into the micropylar opening. In most plant species, the embryo sac is made of seven cells, the two synergids that are just inside the micropylar cavity of the ovule, the egg and central cells and the three antipodal cells. When the pollen tube reaches the micropylar aperture (micropyle) one of the synergids collapses, the pollen tube pass through the filiform apparatus, an invagination created by the cell wall of both the two synergids, and moving inside the degenerated synergid, the pollen tube enters the so called reception phase where it ruptures itself releasing the two sperm cells that are then targeted to the egg and central cell for the double fertilization (Escobar-Restrepo et al., 2007). For a successful fertilization, the reception of the pollen tube inside the embryo sac is crucial and the FERONIA (FER), a RLK (Receptor Like Kinase) encode by the *Arabidopsis thaliana* gene AT3G51550.1 , seems to be

determinant for the correct interaction between the pollen tube and ovule (Haruta et al., 2018b). In *Arabidopsis thaliana*, (fer) mutants, although the plants developed normal shaped female gametophytes, only half of them were fertilized and in those that remained unfertilized, the pollen tube continued to grow, failing to rupture and release of the two sperm cells. Pollen from the same (fer) mutant was able to fertilize wild type (FER) plants (Escobar-Restrepo et al., 2007). A model to describe the two antagonist processes that regulate flower fertilization, pollen tube growth and its disintegration/collapse, was proposed by Escobar-Restrepo et al. (2007) where the interaction between the FER RLK protein and putative ligands produced by the pollen tube, would trigger a signalling transduction cascade into the synergid cells with a subsequent feedback signal that would stop the pollen tube growth and cause the release of the sperms cells. In the mostly 'crumbly' phenotype, the gene corresponding to the *Arabidopsis thaliana* (AT3G51550.1) encoding for the FER RLK is downregulated, with the difference in the expression levels significant at about 95% of confidence level in all the three stages (see Table 3.5), but in particular at the open flower, the only stage where the fertilization process is most likely to take place. This suggests that the mostly 'crumbly' phenotype plants might be operating in a similar way to the *Arabidopsis thaliana* (fer) mutants and partially failing the phase of pollen ligand to RLK interaction, indispensable for the arrest of the pollen tube growth and its release of the sperm cells, with the consequent failed fertilization.

Important compounds like vitamins and hormones belong to one of the largest groups of natural products, the hydrocarbon terpenes, which are the main constituents of plant resins and essential oils (Rohdich et al., 2001). The terpenes derive from the condensation of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) via two biosynthetic pathways, the cytosolic mevalonate (MVA) and plastidic 2-C-methyl-d-erythriol 4-phoshpate (MEP). The first pathway being responsible for the synthesis of sterols and brassinosteroids while the latter for gibberellins, ABA and cytokinins (Suzuki et al., 2009). A key enzyme in the synthesis of IPP and DMAPP via the MVA pathway is the β-hydroxy β-methylglutaryl-CoA reductase (HMGR) since it is considered the rate-limiting factor. In *Arabidopsis thaliana* twelve genes encode for a HMGR and HMG1 and HMG2 have been studied extensively. Both these genes are expressed during male gametogenesis, suggesting that at least one HMGR is required for normal pollen development. The diploid cells of the tapetum express only HMG1 and because those cells are essential for the correct development of the pollen grains, hgm1 homozygous

genotypes are male sterile while by contrast the complete blockage of the MVA pathway does not affect female gametogenesis. One explanation could be that MVA-derived metabolites such as important membrane components like sterols are accumulated in the embryo sac after meiosis (Suzuki et al., 2009). The *Arabidopsis thaliana* (AT1G76490.1) HMG1 raspberry equivalent was downregulated in the mostly 'crumbly' phenotype in all the three different stages analysed, with expression levels significantly different at 95% of confidence level. The biggest difference observed was in the open flower stage; (see paragraph 3.3.3.2, Table 3.16 for details). At the open flower stage, pollen development and maturation take place. Thus, an increase in the expression level of HMGR, compared to the other two stages analysed in this study, could be explained by the contribution of the pollen genes that were expressed only during this stage. It might be the case that in the mostly 'crumbly' plants the significantly lower expression of the HMG1 gene could cause deficient male gametogenesis, producing abnormal gametophytes due to defects in the membrane systems that depend on sterols whose biosynthesis in turn depends on the activity of HMGR enzymes. In *Arabidopsis thaliana*, the tapetum cells express only HMGR1 and this enzyme, at least in this tissue, seems to be indispensable for the MVA pathway and the biosynthesis of cytoplasmic phytosterols (Suzuki et al., 2009). The tapetum cells are essential for the normal pollen formation, a reduction in the expression level of the gene codifying HMGR1, as found in the mostly 'crumbly' phenotype could cause the formation of abnormal pollen due to scarcity of essential membrane components (i.e. phytosterols) whose biosynthesis, at least in the tapetum cell, depends exclusively on HMGR1. Such plants would then produce abnormal pollen, unable to fertilize the ovules. These pollen grains would compete with the functional pollen. In raspberry most varieties being self-fertile can self or cross pollinate, thus in the mostly 'crumbly' plants only the stigmas receiving pollen from normal plants would be able to fertilize the ovules and give rise to drupelets. In the mostly 'crumbly' plants a combination of self and cross pollination would be responsible for the formation of misshapen 'crumbly' like fruits because the abnormal pollen from the mostly 'crumbly' plants would not be able to fertilize the ovules. This model could explain the different degrees at which the 'crumbly' fruit phenotype expresses within and between plants and across seasons where in some cases only a few drupelets form a fruit.

In flowering plants, male gametogenesis begins with the formation of the microsporocyte, a diploid cell that then is subjected to meiosis ending up with the formation of four haploid microspores joined together to form a tetrad by means of a callose cell wall. The

subsequent dissolution of the callose cell wall via hydrolytic enzymes supplied by the sporophytic tapetum allow the release of each individual microspore. Then two mitotic events are required for the maturation of the microspore into a gametophyte. Pollen mitosis I (PM I) is an asymmetric division forming a bicellular pollen grain containing the vegetative cell (VC) and the generative cell (GC). The VC is much larger in size being metabolically active to support the further development of the pollen grain including the production of intine, a layer composed primarily of pectin and cellulose that together with the exine layer forms a cell wall that confers durability to the pollen grain (Kang et al., 2003). In general, the GC is subjected to a second mitosis to form a tricellular pollen grain in which a sperm cell and vegetative nucleus stay associated to form the male germ unit (MGU) that seems to be essential for pollen germination. In *Arabidopsis thaliana*, the gene AT1G14830.1 encodes ADL1C, a member of the family dynamin-like proteins, that seems to be required for the late stages of pollen development. Dynamins are GTP-binding proteins with a GTP binding domain and a GTPase effector domain (Kang et al., 2003). ADL1C seems to be essential for the formation of viable pollen because the *adl1c* mutant in *Arabidopsis thaliana* produce pollen with defects in the intine layer as well as in the plasma membrane of the pollen grains that cannot survive during desiccation, becoming nonviable (Kang et al., 2003). The downregulation of the raspberry gene, equivalent to AT1G14830.1, in the mostly 'crumbly' phenotype, could explain the formation of misshapen 'crumbly'-like fruit since such plants produce more abnormal pollen unable to survive the desiccation phase and become nonviable. The resulting fruit would be misshapen with few drupelets as a result of the limited fertilization by viable pollen.

The *Arabidopsis thaliana* gene AT3G12110.1, encodes for the reproductive actin ACT11. The actin cytoskeleton is a dynamic network composed of actin polymers (actin isovariants) and their binding proteins (ABPs actin binding proteins). The actin cytoskeleton is involved in different physiological processes and within these, the polarized cell growth. *Arabidopsis thaliana* expresses eight functional actin genes that are divided in two groups, the vegetative actins (ACT2, ACT7 and ACT8) and the reproductive actins (ACT1, ACT3, ACT4, ACT11 and ACT12) (Chang and Huang, 2015). Pollen, during its germination and subsequent tube growth, becomes highly dependent on the actin cytoskeleton and after the pollen hydration, endures radical changes. In theory all five reproductive actins' are expected to play a part in the cytoskeleton rearrangement. ACT11 seemed to play an important role in pollen

germination and pollen tube elongation, loss of ACT11 reduces the amount of total actin in the cells with a consequent reduction of filamentous actin (F-actin) levels as well as of its actin isovariants components. These altered actin filaments, compared to those with normal ACT11 level, might interact with different actin binding proteins (ABP) resulting in more dynamic actin cytoskeleton with a consequent negative impact on both pollen germination and tube growth elongation since for instance, in act11 mutants, the actin filaments do not converge properly at the germination aperture causing delay of pollen germination. In fact actin filaments are thought to form the molecular track, responsible for the transport of material indispensable for both membrane expansion and cell wall synthesis that allow the protrusion of the pollen tube (Chang and Huang, 2015). The raspberry gene corresponding to the *Arabidopsis thaliana* AT3G12110.1, was downregulated in the mostly 'crumbly' phenotype with difference being significant at 95% of confidence level in all the three stages analysed (Table 3.13), but with the biggest difference (≈ 10 fold less compared to the never 'crumbly') in the open flower stage.

Auxins are hormones playing important roles across the whole life cycle of plants. The Auxin Response Factors (ARFs) are transcription factors that regulate the expression levels of the auxin mediated genes. Gametophytes are indispensable for plant reproduction, for instance the embryo sac (female gametophyte) is crucial for pollen tube guidance and seed development while the pollen grain (male gametophyte) is crucial for the generation of the sperm cells and for their delivery to the embryo sac. Auxins are involved in the gametophyte development processes and the role of five auxin receptor factors: ARF2, ARF3, ARF4, ARF6 and ARF8 was studied in Liu et al. (2018). The study of transgenic mutants and/or double mutants of all the five ARF transcription factors showed that arf2 mutants had reduced self-fertility due to failure in the opening of floral buds. Single arf3 mutants developed narrow gynoecia, small ovules and aberrant stamen primordia. Double arf3 and arf4 transgenic mutants produced flowers with smaller gynoecia and stamen, while arf5 mutants produces flowers with sterile female and male parts. Single arf6 and arf8 mutants had delayed stamen development and reduced fecundity while the double mutants showed a complete block of the flowering (Liu et al., 2018). In the mostly 'crumbly' phenotypes, the gene corresponding to the *Arabidopsis thaliana* AT1G30330.2 was downregulated. This gene encodes ARF6, that as showed in Liu et al., (2018) is responsible for the correct maturation of stamens and fertility. It might be the case that mostly 'crumbly' type plants behave like *Arabidopsis thaliana* arf6

mutants. The only early visible sign that a plant may be 'crumbly' is a slight browning of the stamens.

The process of formation of plant female gametogenesis, consists of three consecutive mitosis that affect one of the four nuclei of the meiotic spores. The product of these cellular divisions is a megagametophyte, a structure with seven cells and eight nuclei, three antipodal cells, an egg cell, a binucleate central cell and two synergid cells. The megagametophyte development process involves different genes, the FIS-class encoding protein of the polycomb group (PcG) which includes MEA, FIE, FIS2 and the WD-40 repeat protein MSI1 which regulate cell proliferation (Ebel et al., 2004). The retinoblastoma protein (pRB), interacting with specific subunit of PcG and with MSI1, acts as a one of the most important negative regulator of cell division by repressing E2F transcription factor during female gametogenesis (Ebel et al., 2004). Analysis of rbr1 mutants showed that at ovule maturity stage the megagametophyte presents supernumerary nuclei at the micropyle where normally the egg apparatus should be formed. As a result, fertilised rbr1 megagametophytes rarely form embryos and in any case their development is arrested at the globular stage (Ebel et al., 2004). In the microarray experiment, the gene, equivalent to *Arabidopsis thaliana* AT3G12280.1 gene encoding for the retinoblastoma protein (RBR1), was downregulated in the mostly 'crumbly' phenotype, in particular at the open flower stage, suggesting that in those plants, due to aberrant expression of the gene encoding the RBR1, the megagametophytes do not develop well after ovule maturity stage, no egg cell and central cell form, resulting in the formation of an ovary with a disrupted ovule that cannot be fertilised.

The probe CUST_4396_PI426541283 matched a predicted *Arabidopsis thaliana* gene (AT1G42470.1) encoding an ortholog putative Niemann-Pick C1 protein (AtNPC1) which belongs to a class of protein containing a sterol-sensing domain (SSD) and could be involved in the regulation of sterol pathway (Feldman et al., 2015). In *Arabidopsis thaliana*, gene knock out experiments producing homozygous T-DNA insertion lines for (AT1G42470.1 and AT4G38350.1), putative plant ortholog genes of human NCP1, gave evidence of reproductive deficiencies. Such lines were crossed to assess mutant allele segregation in three independent F_2 populations and in two independent F_3 populations and the resulting aberrant segregation gave proof of reproductive defects; furthermore microscopy observation of pollen from these populations showed that less than 50% of pollen was viable compared to 95% of wild plants indicating severe defects in male gametogenesis (Feldman et al., 2015). In the 'crumbly' microarray experiment,

CUST_4396_PI426541283, the probe matching the predicted *Arabidopsis thaliana* gene AT1G42470.1 encoding for a putative ATCNP1 protein, was highly downregulated in the mostly 'crumbly' phenotype. Such low level of expression might simulate the scenario happening in the *Arabidopsis thaliana* T-DNA mutants and so the mostly 'crumbly' plants might produce less viable pollen. Hypothesizing that self-pollination, the one using pollen from the same flower (autogamy) or from the same plant (geitonogamy), could be the most probable event to occur, then in the mostly 'crumbly' plants, the chances for the stigmas to get pollinated by unviable pollen from the same flower/plant would be much higher, simply because these plants produce much more unviable pollen. The pollination with defective pollen negatively affects the fertilization of the ovules causing the formation of fruit with reduced number of drupelets like in the 'crumbly' scenario.

Plant cells are surrounded by the cell wall, a rigid structure composed of different carbohydrate polymers and glycoproteins. Cell growth and expansion takes place only if is accompanied by an increase in cell wall (Pieslinger et al., 2010). Three genes AT3G01640.1 and AT5G15650.1, selected from the analysis of the gene expression in the 'crumbly' microarray experiment, encode enzyme that are directly or indirectly only involved in the biosynthesis of cell wall components.

AT3G01640.1 encodes a GLUCURONOKINASE G, an enzyme localised in the cytosol and responsible for the phosphorylation of phosphorylate D-GlcA to D-GlcA-1-phosphate. The enzyme is involved in the synthesis of UDP-glucuronate, (UDP: uridine diphosphate) a nucleotide sugar involved in the biosynthesis of cell wall components such as carbohydrate polymers and glycoproteins. Glucuronokinase G belongs to the GHMP-kinase superfamily having a unique substrate specificity for d-glucuronic acid. In *Arabidopsis thaliana* the gene is expressed in all plant tissues with a preference for pollen where supplies the cell wall polymers indispensable for accompanying the expansion of the pollen tube (Pieslinger et al., 2010). The second gene AT1G63180.1 encodes the UDP-D-glucose 4-epimerase (UGE3) that catalyses the reversible conversion of UDP-galactose to UDP-glucose (UDP-glc). The enzyme is involved in pollen development (source TAIR https://ui.arabidopsis.org/) and is expressed preferentially in pollen and contributes indirectly to supply the required components to the cell wall during the rapid expansion of the pollen tube (Pieslinger et al., 2010).

Fast growth such as that experienced by the pollen tube during its protrusion inside the style, not only require new component to sustain the expansion of the cell wall but is indispensable a system to deliver these component from their site of synthesis (i.e. cytosol and Golgi apparatus, endoplasmic reticulum) to the target site (cell wall). The two genes, AT3G56950.1 and AT2G18840.1, again differently expressed between the two phenotypes (i.e. mostly and never crumbly), seems to play this role.

AT3G56950.1 encodes a protein belonging to a subfamily of aquaporins, the Small and basic Intrinsic Protein 2 (SIP2;1) involved in the passive transport of small molecules across membranes. In Arabidopsis thaliana (sip2;1) mutants showed normal vegetative growth but presented reproductive anomalies (Sato e Maeshima 2019). The pollen, normal in shape, could adhere to the stigma but presented lower germination rate and limited elongation properties of the pollen tube. SIP2;1 is localised in the endoplasmic reticulum (ER) where is evenly distributed. This cell organelle synthesizes proteins and other molecules for the plasma membrane (PM). It generates small intracellular membrane vesicles that contain the synthesised PM components and the cell wall precursors. These membrane vesicles deliver the components to the PM from the ER through the Golgi apparatus (Sato and Maeshima, 2019). It might be speculated that plant mutants (sip2:1) may have a different distribution of SIP2;1 on the ER and that this feature might reduce the ability of the cells, especially in the fast growing pollen tube, to supply enough cell wall components required to support the fast growth of the tissues. As consequence in the mutants (sip2;1) the pollen tube would not elongate enough to reach the ovule with consequent missed fertilization. AT3G56950.1 was upregulated in the mostly ‘crumbly’ phenotype and that the overexpression of SIP2;1 in mostly ‘crumbly’ plants might be responsible for alteration in the described delivery system of cell wall components that can negatively affect cell elongation in the pollen tube. A consequence of such a scenario might be the reduced rate of fertilization because the pollen tube, for the reasons described above, would stop its growth before reaching the ovule.

The second gene, AT2G18840.1, encodes the YPT/RAB GTPase Interacting Protein 4a (YIP4a) which, on the Golgi apparatus, form a complex with ECHIDNA (ECH), the trans-Golgi network (TGN). This complex is required for the secretion of cell wall polysaccharides that are particularly important to sustain the growth of the cell during elongation (Gendre et al., 2013). One of the components of the TGN complex is (YIP4a), encoded by AT2G18840.1, the gene in the mostly ‘crumbly’ plants was downregulated in all the three development stages (i.e. closed bud open flower and green berry) with difference being statistically significative at 95% confidence levels (see Table 3.25). It

was reasonable to speculate that the under-expression of these gene might affect pollen tube elongation compromising flower fertilisation with consequent formation of fruit with lower number of drupelets that might be misshapen like the ‘crumbly’ one.

3.5 Conclusions

In summary, the microarray experiment on mostly and never 'crumbly' material across three developmental stages (i.e. closed bud, open flower and green berry) allowed the identification of eleven key genes with clear roles in leading ovule fertilization and fruit development. Starting from the list of genes, of the microarray experiment, that were significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly'), screening was done selecting only genes with ontology annotation related to flower development, hormones, pollen and transport. Flower and pollen genes were chosen as during these stages it might be expected that any alterations in the growth processes here could lead to the formation of misshapen 'crumbly' like fruit. The other two groups of genes (i.e. hormones and transport) were considered for their role as plants growth regulators, and transport because they focus on the hormonal interplay between receptacle and fertilised ovaries, proposed in this work (see chapter 1), to regulate the fruit growth and development.

The first screening identified 107 differentially expressed probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* equivalent, according to ontology annotations, were distributed between the classes: (12) flower development (56) hormones, (21) pollen and 20 transport (see Figure 3.13); two genes were shared in two different groups.

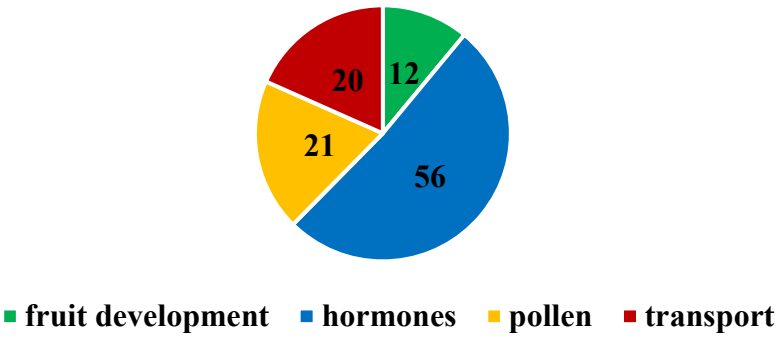


Figure 3.10: Pie chart of the 107 genes differentially expressed between the two different phenotypes (i.e. mostly and never 'crumbly').

From the original list of 827 microarray probes matching *Arabidopsis thaliana* equivalent genes, 107 genes were selected from the four groups (i.e. flower development, hormones, pollen and transport) accordingly to their gene ontology annotations. The majority of the genes 56 (≈ 51%) had gene ontology terms related to hormones, 21 (≈ 19)% related to pollen, 20 (≈ 18%) to transport and 12 (11%) to flower development. Two genes were shared in two different groups.

Three groups (i.e. flower development, pollen and transport) of genes were individually subjected to tree cluster heatmap analysis while the group containing ontology terms related to hormones was further divided in three subgroups (i.e. response to hormones, hormones biosynthesis and 'other'). Each subgroup was analysed separately but always using the same protocol described in section 3.2.3 of this chapter. The tree clustering heatmap analysis allowed further screening of the list of potentially interesting genes linked to 'crumbly' fruit to 69.

From these 69 *Arabidopsis thaliana* equivalent genes, identified via tree clustering heatmap analysis, a further screening was performed by searching, for each gene, for useful information (i.e. description, function and relevant scientific publication) in The *Arabidopsis* Information Resource (TAIR) https://www.arabidopsis.org/ database. The information was considered if it had relevance to any potential mechanism/process linked to the formation of flower and/or fruit, that could reasonably be associated to 'crumbly' like misshapen fruits (e.g. flower differentiation, gametogenesis, fertilization, etc.). The result of this further selection was a list of eleven genes whose products were involved, more or less directly, in processes responsible for pollen formation, pollen maturation, pollen tube elongation, pollen tube recognition and eruption as well as flower development and in particular stamens and ovaries.

The molecular mechanisms taking place at open flower stage and that might, potentially, cause the formation of misshapen 'crumbly' like fruit could be involved in:

- (AT3G01640.1 and AT1G63180.1) - altered synthesis of cell wall component, especially in pollen tube during its protrusion inside the stigma with potential premature arrest of pollen tube elongation and consequent missed fertilization
- (AT3G56950.1 and AT2G18840.1) - impaired transport of cell wall components, respectively from endoplasmic reticulum (ER) and Golgi apparatus, especially in pollen tube during its protrusion inside the stigma with potential premature arrest of pollen tube elongation and consequent missed fertilization
- (AT3G51550.1) - compromised interaction/recognition between pollen tube and embryo sac with consequent lack of fertilization of egg cell by sperm cell
- (AT3G12110.1) - delayed pollen germination due to alteration of the mechanisms involved in cytoskeleton rearrangement important for both pollen germination and pollen tube elongation.

The molecular processes taking place at an earlier stage (closed bud) and that might be responsible for the formation of 'crumbly' fruit might be:

- (AT4G24210.1) - impaired modulation of DELLA proteins (gibberellins (GAs) repressors at transcriptional level) with consequent GAs disequilibrium affecting normal flower formation and pollen maturation
- (AT3G12280.1) - impaired mega-gametogenesis causing the formation of an embryo sac without egg and central cells
- (AT1G76490.1) - formation of unviable pollen lacking essential components of cell membrane systems
- (AT1G14830.1) - altered formation of intine layer (the internal layer of the pollen grain wall) and plasma membranes with consequent formation of flawed pollen
- (AT1G42470.1) - impaired sterol biosynthesis with consequent production of poor viable pollen

The majority of these eleven key genes, from a 'crumbly' fruit perspective, are related to pollen and in particular to its formation and on its functioning. Impaired molecular mechanisms in which those genes are involved might cause formation of poor viable pollen and pollen tube with compromised ability to grow enough into the style and then fertilised the ovule. The final and overall consequence of these impaired processes could be the formation of fruit with lower number of drupelets that might turn to be misshapen and with no commercial value. A scheme summarizing the functions of the six most representative putative 'crumbly' genes was reported in Table 3.27.

closed bud	open flower
AT4G24210.1 encodes SLEEPY1 (SLY1) encoded that forms the complex $SCF^{SLY\text{-}SUMO}$. Such complex modulates DELLA proteins with consequent reduction in gibberellins (GAs) response. Impaired expression of SLY1 can compromise pollen maturation. **Downregulated in the mostly 'crumbly' plants.**	**AT3G01640.1** encodes a **GLUCURONOKINASE G,** expressed in pollen and involved in the biosynthesis of UDP-glucuronate, a precursor of cell wall components. Impairments in its expression might affect the synthesis of cell wall components resulting in pollen tube growth arrest and missed fertilization. **Upregulated in the mostly 'crumbly' plants.**
AT1G76490.1 encodes β-hydroxy β-methylglutaryl-CoA reductase (HGMR1) a key enzyme in the cytosolic mevalonate (MVA) pathway. The gene (HMG1) is expressed in pollen and impairments in its expression can cause deficient male gametogenesis and formation of flawed pollen. **Downregulated in the mostly 'crumbly' plants.**	**AT3G51550.1** encodes FERONIA (**FER**), a receptor like kinase that seems to be determinant for the correct interaction pollen tube-ovule. Impaired expression of this gene might compromise the interaction between pollen tube and FER that is indispensable for the arrest of pollen tube growth and indirectly of ovule fertilization. **Downregulated in the mostly 'crumbly' plants.**
AT1G42470.1 encodes Niemann-Pick C1 protein (**NPC1**), a class of protein containing a sterol-sensing domain (SSD) potentially involved in the regulation of sterol pathway. Anomalous expression of this gene might cause the formation of poor viable pollen. **Downregulated in the mostly 'crumbly' plants.**	**AT3G21110.1** encodes the reproductive actin (**ACT11**), one of the components of the cytoskeleton. It is expressed in pollen and is involved in the rearrangement of the cytoskeleton. Impairments in its expression might cause anomalies in the assemblage of the molecular track that allows the transfer of materials indispensable for both membrane expansion and cell wall synthesis, particularly during pollen tube protrusion. **Downregulated in the mostly 'crumbly' plants.**

Table 3.27: Summary scheme of the impaired key processes taking place in plants showing 'crumbly' fruit symptoms.

Scheme indicating the different impaired processes taking place during close bud stage and open flower stage whose effect might cause the formation of 'crumbly' like misshapen fruit. Irregular expression levels of the three genes on the left might be responsible for anomalies in the pollen formation that are obviously more reasonable to happen at close bud stage. The impaired expression of the three genes on the right side of the scheme might be responsible for anomalies in the activity of pollen and specifically of the pollen tube (e.g. premature arrest of pollen tube growth in the style, pollen tube eruption and release of sperm cells); obviously these processes are more reasonable to occur at open flower stage.

Chapter 4: Using QTL mapping combined with microarray analysis to improve and further understand the genetic control of ‘crumbly’ fruit in the Glen Moy x Latham population

4.1 Introduction

Understanding genetics is key to improving our knowledge of many aspects of plant biology and has advanced greatly through an understanding of the principles of heredity, with Mendelian genetics forming the basis of plant breeding. Together with an understanding of biometrical genetics which applies to traits showing continuous variation and controlled by more than one gene, this has allowed significant understanding of quantitative traits (Mackay, 2001, Rajon and Plotkin, 2013). Using this knowledge, understanding plant trait control has been made possible by the development of genetic linkage maps where markers linked to the gene(s) or quantitative trait loci (QTLs) underlying a trait can be identified.

For map construction, individual marker loci are identified and genetically characterized in a segregating population from a single cross and the recombination rate of alleles at each pair of loci can be determined using classical linkage analysis. These marker loci represent DNA polymorphisms between different individuals. These polymorphisms can be of many different types from single nucleotide changes, large or small insertions and deletions or length variation in repeat sequences.

The first type of DNA marker applied in plant genotyping was the restriction fragment length polymorphism (RFLP); useful in developing genetic linkage maps but limited by complicated hybridization, radioactivity hazard, time consuming procedures and lack of available probes (He et al., 2014). The development of biotechnology techniques and in particular of the polymerase chain reaction (PCR) paved the way for PCR-based markers such as random amplification of polymorphic DNA (RAPD), cleaved amplified polymorphic sequences (CAPS), sequence characterised amplified region (SCAR), and amplified fragment length polymorphisms (AFLP). The great advantages of these PCR-based markers were the relatively lower costs and speed at which they could be identified (He et al., 2014). The limited marker density of some of these methods led to the development of single nucleotide polymorphisms (SNPs) that have become the preferred markers to be used in genotyping and mapping (Beissinger et al., 2013).

SNPs are the most abundant markers in a genome and can be used for many different genome analyses (He et al., 2014). These marker loci can be applied to a biparental cross and ordered into a linkage map and distance between loci can be expressed as recombination units given in centiMorgans (cM) where one cM is equal to 1%

recombination. Once a sufficient number of markers have been mapped, the number of linkage groups (LGs) should equal the haploid number of chromosomes. Computer programs are available to quickly generate a map once markers have been applied to a segregating population (e.g. JoinMap software by Kazyama®).

Compared to other crops, there have been few genetic and genomic resources available for *Rubus* until recently when the development of high-density genetic maps and markers and other genomic resources have become available.

The James Hutton Institute played a role in contributing to the development of new genomic resources, for instance creating a database that provides access to the predicted genes from an assembly of whole-genome-shotgun sequence from raspberry (cv. Glen Moy). The genes were predicted from the mapping of RNA-sequencing data from twenty-two varieties of raspberry to the genomic assembly (http://camel.hutton.ac.uk/raspberry/).

For reviews of markers, linkage maps and QTL developments in *Rubus* see (Graham et al., 2009, McCallum et al., 2018, Foster et al., 2019). Briefly, the first raspberry genetic map using markers was developed by Graham et al., (2004) from a 'Glen Moy' x 'Latham' population and this map has subsequently been improved over a number of years (Graham et al., 2009, Woodhead et al., 2010, Graham et al., 2011, Kassim et al., 2009, McCallum et al., 2010, Paterson et al., 2013, Simpson et al., 2017, Graham et al., 2006) and recently by using Genotype by Sequencing (GBS) (Hackett et al., 2018). Aside from the 'Latham' x 'Glen Moy' population, a number of other linkage maps in red raspberry have been generated. A 'Latham' x 'Titan' map (Pattison et al., 2007), a 'Malling Jewel' x 'Malling Orion' map (Sargent et al., 2007). Ward et al., (2013) used Genotyping by Sequencing (GbS) to produce highly saturated maps for a *R. idaeus* pseudo-testcross progeny. Castro et al. (2013) published the first genetic map of a primocane-fruiting and thornless tetraploid blackberry (*Rubus* subgenus *Rubus* Watson). Bushakra et al., (2015) constructed the first linkage map of black raspberry (*Rubus occidentalis*) using single-nucleotide polymorphism and simple sequence repeat markers representing seven linkage groups. A cross between black and red raspberry was used in a comparative genomic mapping study to align the *Rubus* linkage groups with those of strawberry (Bushakra et al., 2012).

The developments of GbS mapping has allowed high density maps linked to genome scaffolds to be available for trait dissection (Foster et al. 2019). The new generation of sequencing technologies allows high-throughput at lower costs; moreover, due to their

accuracy and simplicity they are becoming commonly used in both bi-parental mapping and genome wide association studies (GWAS). The GbS technology is quite straightforward when it comes to small genomes while for large ones, enrichment or restriction strategies must be applied to guarantee enough overlap of sequence coverage. While enrichment strategies are time consuming and extremely expensive for large genomes, the reduction of the genome complexity with restriction enzymes is easy, site specific, reproducible and facilitate to access genome regions inaccessible to sequence capture techniques. Above all, when appropriate restriction enzymes are selected, repetitive regions of the genome as well as lower copy regions are easily avoided simplifying the computational alignment in species with high level of genetic variation (Elshire et al., 2011).

The use of GbS to identify new SNP markers and construct new linkage maps is becoming increasingly common in genetic studies especially for analysis of quantitative trait loci (QTLs); these are segments of genome that show statistical significant association with quantitative traits (Moose and Mumm, 2008). GbS analysis proved to be successful in the identification of segregation distortion in raspberry, allowing the detection of deleterious alleles responsible for inbreeding depression in *Rubus idaeus* (Ward et al., 2013).

In terms of 'crumbly' fruit, previous QTL mapping of the 'crumbly' fruit trait in the Latham x Glen Moy population revealed two QTLs important for the genetic control of the condition located on linkage group 1 and 3, (LG1) and (LG3) (Graham et al., 2015). Contrary to the suggestion by Jennings (1967b) that 'crumbly' fruit was related to the gene *H* region, no genetic association with this region on LG2 could be identified with the crumbly fruit syndrome (Jennings, 1967). Gene *H* controls cane pubescence (i.e. genotype *Hh* with the dominant genotype *HH* being lethal) while the recessive genotype (*hh*) causes glabrous canes. However, there was an association with ripening, with the longer the fruit takes to achieve fruit set and reach the green fruit stage, the more likely it is to be crumbly. This may explain the association hypothesized by Jennings as the *Hh* genotype of gene *H* is associated with a slowing down of ripening across all stages from open flowers to the green/red stage compared to the *hh* genotype (Graham et al., 2009).

In this chapter data is presented on the genetic basis of the 'crumbly' phenotype using the GbS map of Hackett et al., (2018) to re-analyse the previous data from Graham et al. (2015). The probes of the 'crumbly' microarray experiment (see chapter 3 for details) that mapped inside the 'crumbly' QTLs identified using the GbS map were selected and analysed. The idea was to identify potentially interesting genes related to 'crumbly' fruit

whose relevance would have been strengthen by their location inside a genome region already tightly associated to the 'crumbly' fruit phenotype such as the 'crumbly' QTLs. The aim was to try to find putative gene markers and in fact the new GbS linkage map allowed the identification of new and more significant 'crumbly' markers associated with the 'crumbly' QTLs that could be used for molecular markers assisted breeding and for diagnostic purposes. Moreover, the new GbS linkage map was used by Hackett et al., (2018) to re-analyse the fruit ripening scores (Graham et al., 2009) and identify QTLs associated to the ripening process. In this work, putative association between 'crumbly' fruit and fruit ripening were studied by analysing the functions of those *A. thaliana* genes orthologs of the *R. idaeus'* ones matched by the microarray probes mapped inside the overlapping regions of 'crumbly' and fruit ripening QTLs.

4.2 Materials and methods

4.2.1 'Crumbly' QTLs identification and their mapping on Glen Moy x Latham linkage groups; re-analysis of 'crumbly' data on new GbS map

The crumbly data from the Latham x Glen Moy population (Graham et al., 2015) was re-analysed on the GbS map as described for ripening traits in Hackett et al. (2018) and here data from the 'crumbly' fruit microarray experiment (see chapter three for further details) were integrated in the analysis. The 'crumbly' data QTL positions were identified using a hidden Markov model (HMM) approach adapted from similar work for QTL mapping in autotetraploid species (Hackett et al., 2013), as initial QTL mapping using interval mapping in MapQTL 5 (Van Ooijen, 2004) gave logarithm of the odds (LOD) profiles that were unexpectedly irregular, given the high-density map, resulting in uncertainty in locating the peak LOD score. This was due to the differences in the number of parental heterozygous markers (Hackett pers. Comm.) thus HMM was employed. The hidden Markov model utilising all the marker information on a chromosome allowed the derivation of genotypic probabilities at each position, producing more accurate and precise peaks with smoother profiles (Hackett et al. 2018).

4.2.2 Identification of differentially expressed genes within 'crumbly' QTLs

All probes from the microarray were located onto the new GbS map (Hackett et al., 2018) to look for co-location with the 'crumbly' QTLs. A list with all the probes, mapped inside the three 'crumbly' QTLs, the two identified by Graham et al., (2015) on linkage groups (LGs) 1 and 3 and the new QTL again LG3 identified here, was produced by using the proprietary *Rubus idaeus* genome browser at The James Hutton Institute (http://camel.hutton.ac.uk/raspberry/index.html) to find the probe position within the genome scaffolds. An orientation file that identified the position of genome scaffolds for each linkage group was then utilised. A total of 1,375 probes, 307 from the QTL, identified in this work, on LG3, 851 and 216 probes respectively from the original 'crumbly' QTLs on LG1and LG3 identified by Graham et al., (2015) were identified. The selection of the Glen Moy scaffolds within the 'crumbly' QTLs was based on a 2 cM support interval on either side of the most representative QTL marker. The selected

probes were subjected to analysis of variance (ANOVA) with a phenotype (i.e. mostly and never 'crumbly) per stage (i.e. closed bud, open flower and green berry) interaction design. A very stringent threshold (p-value <0.001) to select the probes differently expressed was applied. The analysis was performed in GenStat (VSN international, UK) to identify those showing differential patterns of expression.

The 165 probes selected as being significantly differently expressed, with regards to stage*phenotype interaction were matched to *Rubus idaeus* genes that were selected from the James Hutton Institute (JHI) Glen Moy data base. The Glen Moy genome browser provides for each gene the sequence that was blasted into The Arabidopsis Information Resource (TAIR) database (https://www.arabidopsis.org/) to identify *A. thaliana* ortholog genes to the selected *R. idaeus'* ones. These selected *Arabidopsis thaliana* genes were analysed with the gene ontology (GO) annotation function of the TAIR database to look for otology terms related to flower development, hormones, pollen and transport as was previously done in chapter three.

4.2.3 Primers design for 'crumbly' marker identification

The new and denser Genotype by Sequencing (GbS) map (Hackett et al., 2018) allowed the identification of new and more significant markers for each of the three 'crumbly' QTLs; the two previously identified by Graham et al., (2015) on linkage groups (LGs) 1 and 3 and the new QTL, identified in this work, again on LG3. These markers represent, potentially, very strong and powerful tools because they could be employed in breeding programme to select new varieties 'crumbly' free and to design diagnostic protocols (i.e. plant health certification scheme).

The primers, both forward and reverse, were designed to the most significant markers in the 'crumbly' QTLs. In total five single nucleotide polymorphisms (SNPs) and three simple sequence repeats (SSRs) were selected and primers designed with exactly the same procedure by using, as reference, the Glen Moy genome database (http://camel.hutton.ac.uk/raspberry/index.html). Three of these eight markers were target genes, located in coding regions while the other five were gene 'tags', located in non-coding regions of the genome (Collard et al., 2005); they were chosen, one for each 'crumbly' QTL, because they mapped in proximity to the most representative marker of the corresponding 'crumbly' QTL.

In Tables 4.13 to 4.16 the information about these three Glen Moy genes were reported. In the Glen Moy genome assembly browser the sequence of the scaffolds, where these genes were located was selected; the software allows the ability to highlight SNPs and here their position was noted together with the position of the genes along its scaffold. The full scaffold sequence was inserted in a specific software, Sequencher® version 5.4.6 (Gene Codes Corporation, Ann Arbor, MI USA). The software displays the full sequence with the position of each base pair (bp) allowing selection of the region of interest. For primer design a portion of the gene sequence (600 bp) containing as many as SNPs possible was selected and inserted into the online software Primer3web version 4.1.0. (primer3.ut.ee). For the other five markers a similar procedure was used, the only difference was that in this case there was no specific gene of reference to look at but only a position along the scaffold where the markers were located. In the Glen Moy assembly database the sequence of each scaffold was copied, always after having noted the flanking regions containing the markers to facilitate their selection in Sequencher® version 5.4.6 the region (600 bp) for use in designing the primers with the software Primer3web version 4.1.0. (primer3.tu.ee). For the SNPs markers the goal was to design primers, amplifying a region containing as many as SNPs possible, while for SSR markers, as long a repeat sequence as possible was identified to increase the chance of the region being polymorphic across accessions.

The designed primers for Sanger sequencing were purchased from Eurofins Genomics (Ebersberg, Germany). The three primers for fragment analysis (genotyping) were purchased from Sigma-Aldrich (USA) and of these, the forward primers were HEX (hexachloro-fluorescein) labelled.

4.2.4 Sanger sequencing for SNPs detection

The fluorescent Sanger sequencing was carried out by the Genome Technology lab at The James Hutton Institute (Dundee, Scotland, UK). DNA was extracted (see section 2.2.2) from all the sixty-three different raspberry cultivars and selections (see Table A.1 appendix). A 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) was prepared containing 0.2 mL of DNA stock solution (1:10 dilution) for each of the sixty-three samples. A Polymerase Chain Reaction (PCR) to amplify the region to be sequenced was performed as followed, 10 µL of DNA stock solution (1:10) were mixed with 15 µL

of PCR reaction mix. The mix prepared for 100 reactions contained 250 µL of dNTPs, Deoxynucleoside triphosphates (Invitrogen™ Corporation, USA), 250 µL of buffer mix and 20 µL of Taq DNA polymerase both (10 µM) and both from (F. Hoffmann-La Roche, Switzerland), 10 µL forward primer and 10 µL reverse primer Eurofins Genomics (Ebersberg, Germany) and 960 µL sterile distilled water. The plate with the PCR reactions was then loaded in a thermocycler, Alfa Thermo Cycler (PCR max, UK) with the following reaction settings (see Table 4.1).

Temperature °C	Time (min)	no. of cycles
95	**5**	**1**
94	**1**	**35**
57		
72		
72	**8**	**1**

Table 4.1: PCR settings for the amplification of the genomic regions to be sequenced for markers detection.

After amplification, the DNA samples were cleaned-up to remove any residual dNTPs that could affect the sequencing reaction; 5 µL of PCR reaction for each DNA sample were transferred to a new 0.2 mL non-skirted 96-well PCR plates and the enzymatic clean-up protocol was performed by adding, 1 µL of Shrimp Alkaline Phosphatase (rSAP) and 1 µL of Exonuclease I both from New England BioLabs Inc (USA) to each sample. The PCR plate was then moved in the thermocycler, for the second and last step of the clean-up protocol, with the following settings, 37 °C for 15 minutes and 80 °C for another 15 minutes; the latter temperature to completely and irreversibly inactivation of the enzymes. Once the clean-up reaction was complete, the samples were sent to the Genome Technology lab at The James Hutton Institute for the Sanger sequencing; the ideal amount of DNA for the sequencing for amplified fragments of size between 200 and 500 bp would be 3-5 ng; an estimation of the amount of DNA was done by running few random samples on a 1.5% agarose gel and then comparing the DNA band intensity with those of known amount of DNA used as reference.

For the sequencing reaction, the 5 µL of samples were further processed by adding 1 µL of (10 µM) of forward primer, 1 µL of Big Dye Terminator (version 3.1) reaction mix Applied Biosystems S.A. (USA), 1.5 µL of 5X dilution buffer Applied Biosystems S.A. (USA) and distilled water to bring final reaction volume to 10 µL. The plate containing the samples was then transferred to TETRAD thermal cycler, Applied Biosystems S.A. (USA), using the following programme, hold at 96 °C for 1 minute and then a cycle to be repeated 25 times at 96 °C for 10 seconds, 50 °C for 5 seconds and 60 °C for 4 minutes. After amplification, the samples were cleaned up by adding for each 10 µL of reaction 2.5 µL of EDTA (125 mM; ph 8.0) and after a brief spinning 30µl of 95% ethanol. The sample were vortexed and spun briefly, incubated at room temperature for 15 minutes and then spun for 30 min at 3000 revolutions per minute (rpm) by keeping the sample at 4 °C. After that, the samples were spun upside down at 100 g for 10 sec, a further cleaning up step was performed by adding 150 µL of 70% ethanol. The sample were briefly vortexed and spun for 10 minutes at 3000 rpm at 4 °C, again this step was followed by spinning the samples upside down at 100 g for 10 seconds; this whole cleaning step was then repeated and after that the samples were left drying at room temperature. Once dried the samples were re-suspended in 10 µL of highly deionized (Hi-Di) formamide, Applied biosystems S.A. (USA) and then analysed on the capillary sequencer ABI3730 DNA analyser Applied biosystem S.A. (USA) with a 36 cm capillary array and Applied Biosystem's POP7 polymer. The samples were run using Applied Biosystems 3730 Data collection software version 4.0 and the sequencing Data coming off the machine were analysed using Applied Biosystem's Sequence Analysis version 6.0; all software Applied Biosystem S.A. (USA).

4.2.4 Fragment analysis for genotyping to detect SSR markers

The DNA SSR genotyping was carried out by the Genome Technology Lab at The James Hutton Institute (Dundee, Scotland, UK). The DNA samples, used for this analysis, were those utilised for the sequencing so the same DNA extraction procedure was applied (see section 2.2.2). The PCR amplification, of the genome regions containing the markers, was performed in the same way described in section 2.3.4; the only difference was the use of fluorescently labelled forward primers. After amplification, 5 µL of the PCR reactions were run on an electrophoretic gel (1.5 % agarose) to estimate the amount of DNA and then, accordingly, to calculate the dilution rate; usually 280 µL of sterile

distilled water (SDW) per PCR reaction were added. To prepare the samples for the analysis, a reaction mix containing 10 µL of Gene Scan™ 500 ROX™ and 990 µL of highly deionized (Hi-Di) formamide both Applied Biosystems S.A. (USA); the 1 mL volume (ROX + Hi-Di) reaction mix refers to 100 reactions. Two µL of SDW diluted PCR reaction were mixed with 8 µL of ROX mix reaction buffer and transferred to a new 0.2 mL non-skirted 96-well PCR plates (ThermoScientific, UK) for the fragment analysis (genotyping). The samples were processed with the Genetic Analyzer 3730 Applied Biosystems S.A. (USA), the data were collected by means of the Genetic Analyzer 3730 data collection software version 4.0 Applied Biosystems S.A. (USA). The final data analysis was performed with GeneMapper software version 5.0 Applied Biosystems™ (USA).

4.2.6 Selection of a population of raspberry genotypes for markers validation pool.

Sixty-three different genotypes (see Table A.4.1 in appendix) were selected to allow the identification of any associations with the phenotype and the potential 'crumbly' markers that could be of economic importance for molecular breeding programs and for diagnostic purposes (i.e. plant health certification). The population of genotypes was carefully selected following consultation with Nikki Jennings (raspberry breeder at The James Hutton Limited), however this selection was not straight forward as some conflicting information was available on the status of the genotypes depending on location and season ~~and this was a noted limitation in the GWAS~~. Forty-five genotypes, a mix of cultivars and selections, both primocane and floricane fruiting forms, were selected as plants that had been reported as showing 'crumbly' symptoms. The remaining eighteen genotypes, circa 25% of the total, were again a mix, of cultivars and selections both floricane and primocane but these were selected because they were never reported to have displayed 'crumbly' symptoms anywhere in any season and this was the reason for a low number of non 'crumbly' genotypes being included. Where possible, the plant pedigree was recorded because one of the criteria for the analysis to identify potential robust 'crumbly' markers, was to test them on a population of unrelated genotypes. Such approach would permit the selection of markers with a broad applicability.

4.2.7 Statistical analysis

All the statistical analyses were performed in GenStat 18th edition (VSN International, UK). The ANOVA with stage (i.e. closed bud, open flower and green berry) per phenotype (i.e. mostly and never 'crumbly') interaction was effectuated for verifying the significative differences, at 99.9% confidence levels, in the expression levels of the 'crumbly' microarray probes mapped inside the three 'crumbly' QTLs. The chi-square analysis and the permutation test were performed to identify statistically significative association between the selected 'crumbly' markers and the population of 63 genotypes loosely related to each other with about ¾ of them labelled as showing 'crumbly' symptoms while the remaining ¼ being never 'crumbly'.

4.3 Results

4.3.1 QTL analysis

The 'crumbly' field scores used here were those recorded over a period of seven fruiting seasons (Graham et al., 2015). The scorings were recorded to investigate the segregation of 'crumbly' fruit in a Latham x glen Moy cross. None of the progenies were scored as 'crumbly' in every season, the maximum was about 75% of the scoring times. However, some individuals never exhibited the 'crumbly' phenotype in any of the seven seasons scored. The Glen Moy x Latham linkage map consisted of 439 markers and allowed the identification of two 'crumbly' QTLs respectively on Linkage Group (LG) 1 and 3 (Graham et al., 2015).

These same 'crumbly' scores were re-analysed in this study using the enhanced high density (GbS) linkage map to which 2348 new high confidence SNP markers were added. This linkage map with higher coverage of markers allowed identification of a new QTL as well as a more accurate genetic position of loci detected with the previous linkage map (Hackett et al., 2018). To identify a total of 2348 new high confidence SNPs, the first draft genome assembly was created, using the cultivar Glen Moy as reference genome. Three hundred and sixty Mbp of genome sequence ordered in 147,546 scaffolds were produced (Hackett et al., 2018). From the 'crumbly' fruit perspective, the results of the analysis of the 'crumbly' scorings with the new GbS linkage map were the confirmation of the two QTLs on LG1 and LG3 previously identified by Graham et al., (2015) with now, the most representative markers being **s3407_p12510_R23 at 7.1** cM (see Figure 4.1) and **s182_p91185_R6** at 106.4 cM (see Figure 4.3) respectively. An interesting finding was the discovery of a new QTL on LG3 not previously identified and having as most representative marker **Rub2a1** at 64.2 cM (see Figure 4.2). All the three markers have four alleles (ab in Latham and cd in Glen Moy) and the significant markers always segregated in the Latham parent (the mean proportion of 'crumbly' fruit was higher in the genotypes carrying the Latham 'b' allele) for the two previously identified QTLs while for the newly identified 'crumbly' QTL on LG3, the allele combination 'bc' gave higher occurrence of 'crumbly' phenotype than the others, suggesting that both alleles were dominant and then the most significant markers segregate in both parents (Table 4.2).

The point estimation of the QTL location, the position along the map with the highest LOD for that specific QTL is the first step in the process of identification of a QTL. It is important to estimate the QTL precision by means of a confidence interval (supporting interval) and here, the 'LOD drop-off method' (Hackett, 2002) was applied (see section 2.6.1).

Year	env.	LG	trait	§pos.	mean				%var	[a]LOD	S.E.				detected previously	key parent
					ac	ad	bc	bd			ac	ad	bc	bd		
2007	field	1	crumbly*	6	0.0325	0.1717	0.44	0.382	10.09	5.05	0.0735	0.0547	0.0693	0.0549	yes	Latham
2009	field	1	crumbly*	7	0.0589	0.1541	0.4517	0.5365	16.62	8.12	0.0728	0.0556	0.0687	0.0562	yes	Latham
2010	field	1	crumbly*	0	0.3444	0.1595	0.4439	0.5468	9.07	4.7	0.0768	0.06	0.0784	0.064	yes	Latham
2011	field	1	crumbly*	9	0.7709	0.7415	1.511	1.5776	14.56	6.45	0.1678	0.1279	0.1673	0.1271	yes	Latham
2012	field	1	crumbly*	9	0.6913	0.6321	1.5116	1.6296	15.91	6.05	0.2079	0.1584	0.1898	0.148	yes	Latham
2007	field	3	crumbly*	59	0.206	0.2014	0.5333	0.1075	11.8	6.05	0.0525	0.0704	0.0602	0.0638	no	[b]both
2009	field	3	crumbly*	59	0.2639	0.2443	0.5952	0.129	12.32	6.31	0.055	0.0737	0.0631	0.0668	no	[b]both
2010	field	3	crumbly*	57	0.2591	0.3517	0.5919	0.1572	10.33	5.25	0.0602	0.0782	0.0614	0.0803	no	[b]both
2011	field	3	crumbly*	107	0.7845	0.948	1.5999	1.7028	13.43	6.04	0.1227	0.1382	0.1502	0.1868	yes	Latham
2011	field	3	crumbly[ł]	107	0.7845	0.948	1.5999	1.7028	13.43	6.04	0.1227	0.1382	0.1502	0.1868	yes	Latham
2012	field	3	crumbly[ł]	54	0.5463	0.5677	0.8801	0.4645	10.19	4.44	0.0681	0.0877	0.0671	0.0906	no	[b]both
2012	field	3	crumbly[ł]	106	0.7324	0.9398	1.627	1.6786	11.96	4.75	0.1512	0.1712	0.1795	0.2165	yes	Latham

*incidence scores
[ł]severity scores
§pos.(position) it refers to the position in cM of the marker with the highest LOD
[a]LOD logarithm of the odds
[b]Latham and Glen Moy

Table 4.2: 'Crumbly' QTLs detected using a Hidden Markov Model (HMM).

'Crumbly' QTLs detected using the Hidden Markov Model (HMM). Env. (environment), LG (linkage group), pos. (position of the maximum LOD in cM), mean 'ac' etc. (the means of the four offspring genotype classes: ac, ad, bc and bd assuming a QTL model of Latham with genotype ab and Glen Moy with genotype cd), S.E. 'ac' etc. (corresponding standard errors of means), %var (the percentage trait variance explained by the QTL), detected previously (refers to whether the QTL was reported by Graham et al. (2015) and key parent (indicates whether one or both parents' alleles were significant). The severity scores referred to a 0-4 scale of 'crumbliness with 0 (no 'crumbly') and 4 (severe 'crumbly' condition).

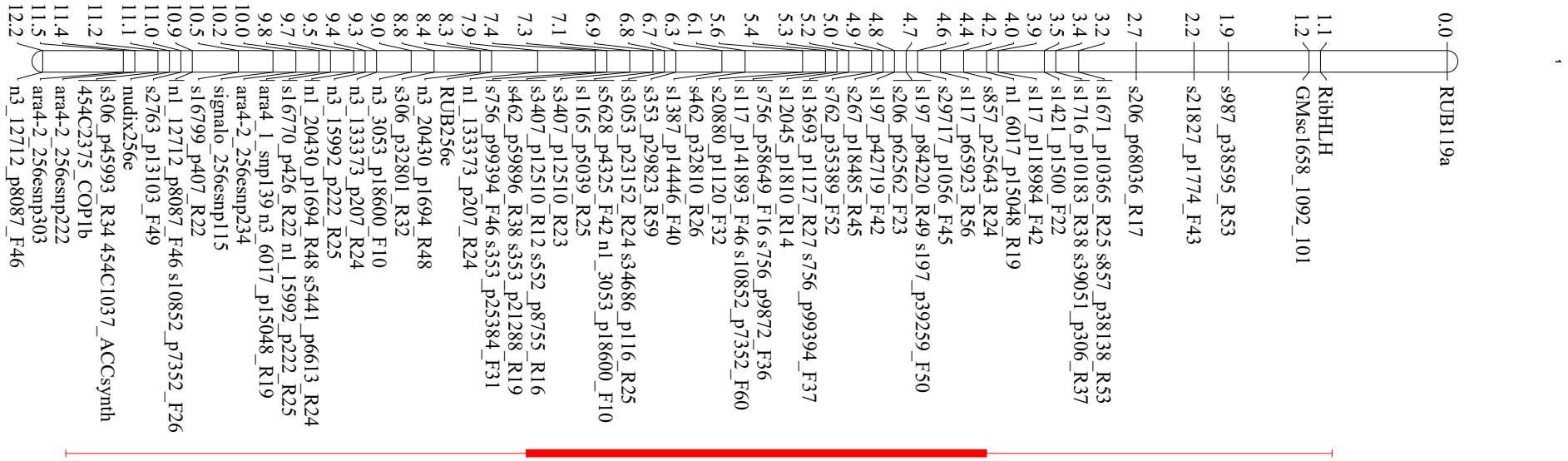


Figure 4.1: Genotype by Sequencing partial map for linkage group 1.

Genotype by Sequencing (GbS) partial map for the linkage group one (LG1) for the Glen Moy x Latham mapping population. For linkage group 1, the most robust marker (**s3407_p12510_R23**) of the 'crumbly' QTL being mapped at 8.3 cM, The QTL support interval (i.e. 1-12 cM) defined by the 'LOD drop-off was represented by a red bar with whiskers.

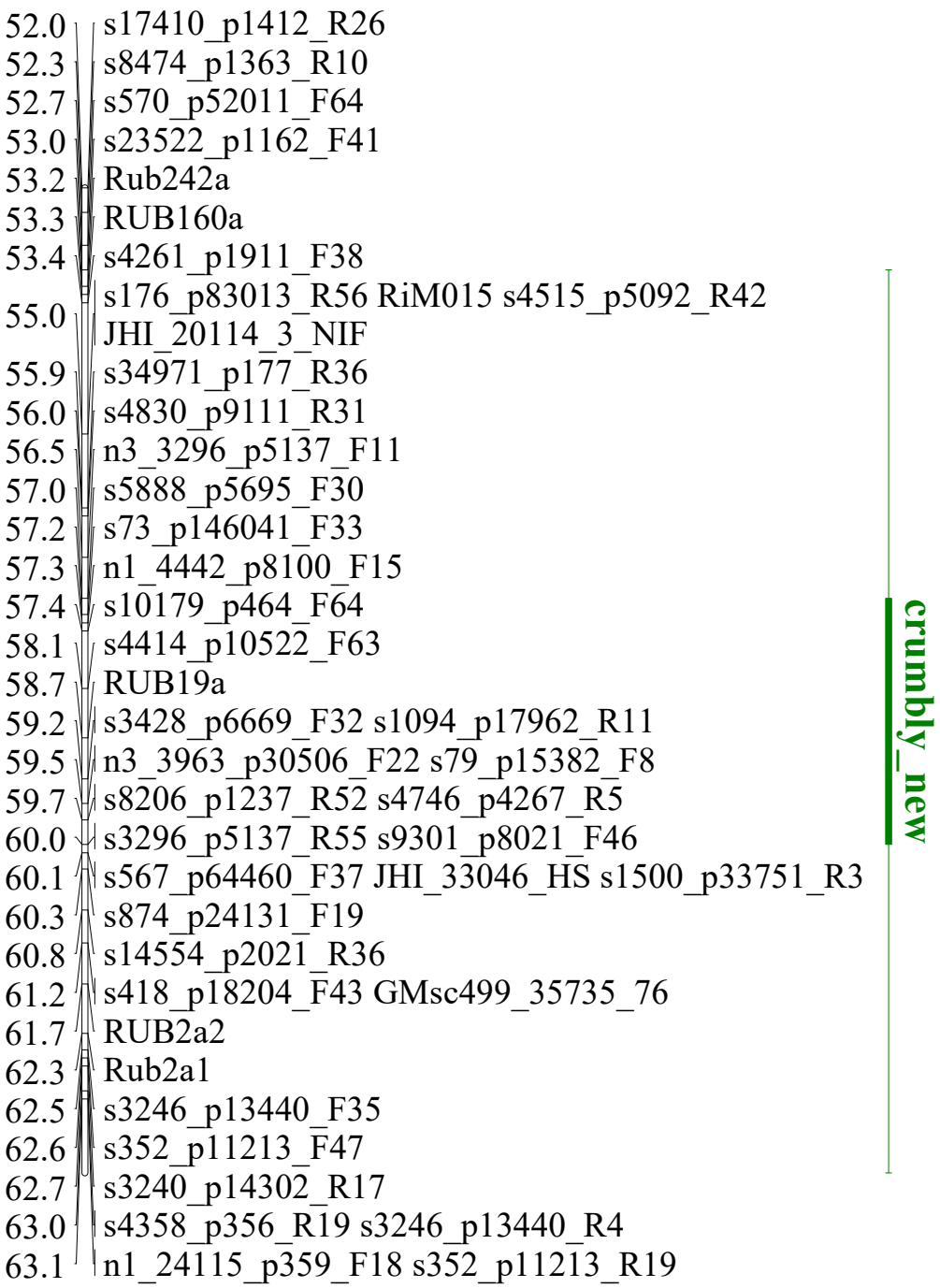


Figure 4.2: Genotype by Sequencing partial map for linkage group 3.

Genotype by Sequencing (GbS) partial map for the linkage group three (LG3) for the Glen Moy x Latham mapping population. The Figure shows the new 'crumbly' QTL, here identified during this work on Linkage group 3. The most robust marker (**Rub2a1**) being mapped at 62.3 cM. The QTL support interval (i.e. 53-62 cM) defined by the 'LOD drop-off was represented by a green bar with whiskers.

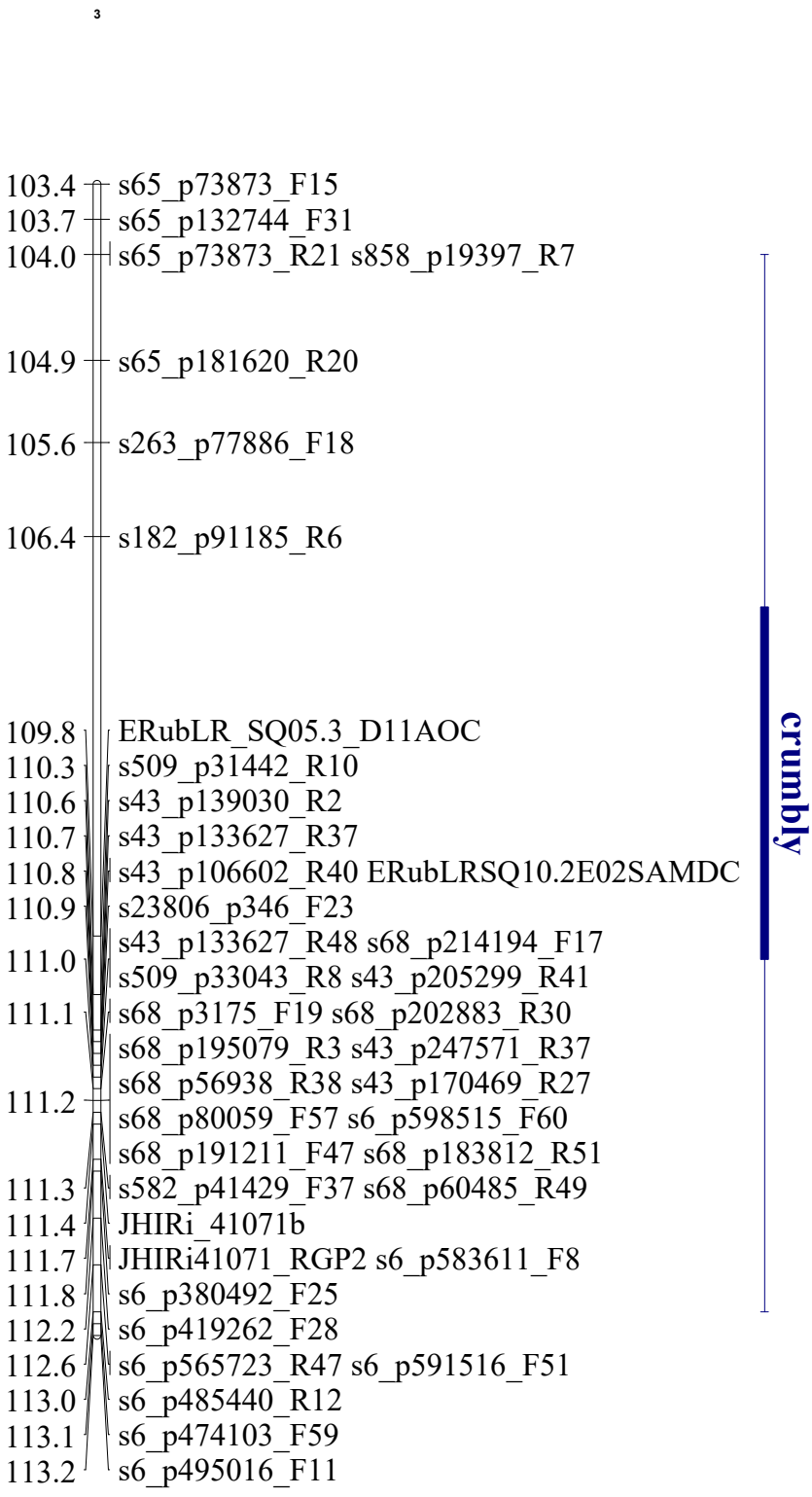


Figure 4.3: Genotype by Sequencing partial map for linkage group 3.

Genotype by Sequencing (GbS) partial map for the linkage group three (LG3) for the Glen Moy x Latham mapping population. The Figure shows the second 'crumbly' QTL for linkage group 3. This QTL was identified by Graham et al., (2015) and has as most robust marker (**s182_p91185_R6**) being mapped at 106.4 cM. The QTL support interval (i.e. 104-112 cM) defined by the 'LOD drop-off was represented by a blue bar with whiskers.

4.3.2 Microarray analysis and probes location on crumbly QTLs

All the 'crumbly' microarray probes on scaffolds within the QTLs were identified. These were analysed by means of the *Rubus* genome browser, to select all those probes matching the scaffolds linked to the three crumbly QTLs. Scaffolds within the QTLs are given in Table 4.3. All these probes were subjected to analysis of the variance (see section 4.2.7) to select those significantly differentially expressed between mostly and never 'crumbly' fruits, after data pre-treatment and normalization (see chapter three for more details).

Linkage group 1		Linkage group 3		Linkage group 3		
QTL	scaffold267	QTL	scaffold6910	QTL recently identified during this work	scaffold73	scaffold1499
	scaffold762		scaffold10689		scaffold4442	scaffold4488
	scaffold12045		scaffold664		scaffold10179	scaffold1348
	scaffold756		scaffold676		scaffold880	scaffold4923
	scaffold117		scaffold835		scaffold6039	scaffold2011
	scaffold10852		scaffold936		scaffold1541	scaffold2770
	scaffold20880		scaffold1200		scaffold1514	scaffold1348
	scaffold462		scaffold3183		scaffold5422	scaffold2505
	scaffold1387		scaffold826		scaffold3486	scaffold1076
	scaffold353		scaffold1925		scaffold3264	scaffold1814
	scaffold3053		scaffold2858		scaffold9301	scaffold3006
	scaffold34686		scaffold4283		scaffold5888	scaffold1868
	scaffold5628		scaffold5185		scaffold2398	scaffold3963
	scaffold1165		scaffold76		scaffold6951	scaffold1348
	scaffold3407		scaffold162		scaffold674	scaffold8206
	scaffold552		scaffold52		scaffold1340	scaffold4746
	scaffold2164		scaffold294		scaffold4414	scaffold3296
	scaffold20430		scaffold4283		scaffold4005	scaffold9301
	scaffold306		scaffold734		scaffold1340	scaffold567
	scaffold318		scaffold65		scaffold44367	scaffold1500
	scaffold15992		scaffold43		scaffold419	scaffold874
	scaffold5441		scaffold858		scaffold879	scaffold2342
	scaffold16770		scaffold263		scaffold79	scaffold1891
	scaffold16799		scaffold182		scaffold6317	scaffold4425
	scaffold12712		scaffold509		scaffold2505	scaffold14554
	scaffold2763		scaffold43		scaffold2770	scaffold418
	scaffold4961				scaffold3428	scaffold499
	scaffold260				scaffold1094	scaffold1868
	scaffold31304				scaffold2954	scaffold698
	scaffold2451				scaffold377	scaffold2268

Table 4.3: Genome scaffolds (in map order) within the three crumbly QTLs, the one on linkage group 1 and the two on linkage group 3.

4.3.3 ANOVA of microarray probes mapped within the three crumbly QTLs

The probes within the scaffolds in the three QTLs were subjected to a general analysis of variance (ANOVA) with a phenotype (i.e. mostly and never 'crumbly) per stage (i.e. closed bud, open flower and green berry) interaction design to confirm those differently expressed between the two different phenotypes.

Of the 1,273 probes identified within the three QTLs, 276 were within the QTL on LG1, 307 inside the new QTL on LG3 and 710 within the previous identified QTL also on LG3. Of these 1,273 probes, only 165 were differently expressed with differences being statistically significant at 99.9% confidence level (p-value <0.001). The probes were distributed thus, 21 on LG1, 38 on the new QTL on LG3 and 106 on the previously identified QTL on LG3.

In Figures 4.4 and 4.5 were reported the tree cluster heat maps for all the differently expressed probes mapped respectively within the 'crumbly' QTLs of LG1 and LG3. All the statistical analyses were performed in GenStat 18th edition (VSN International, UK).

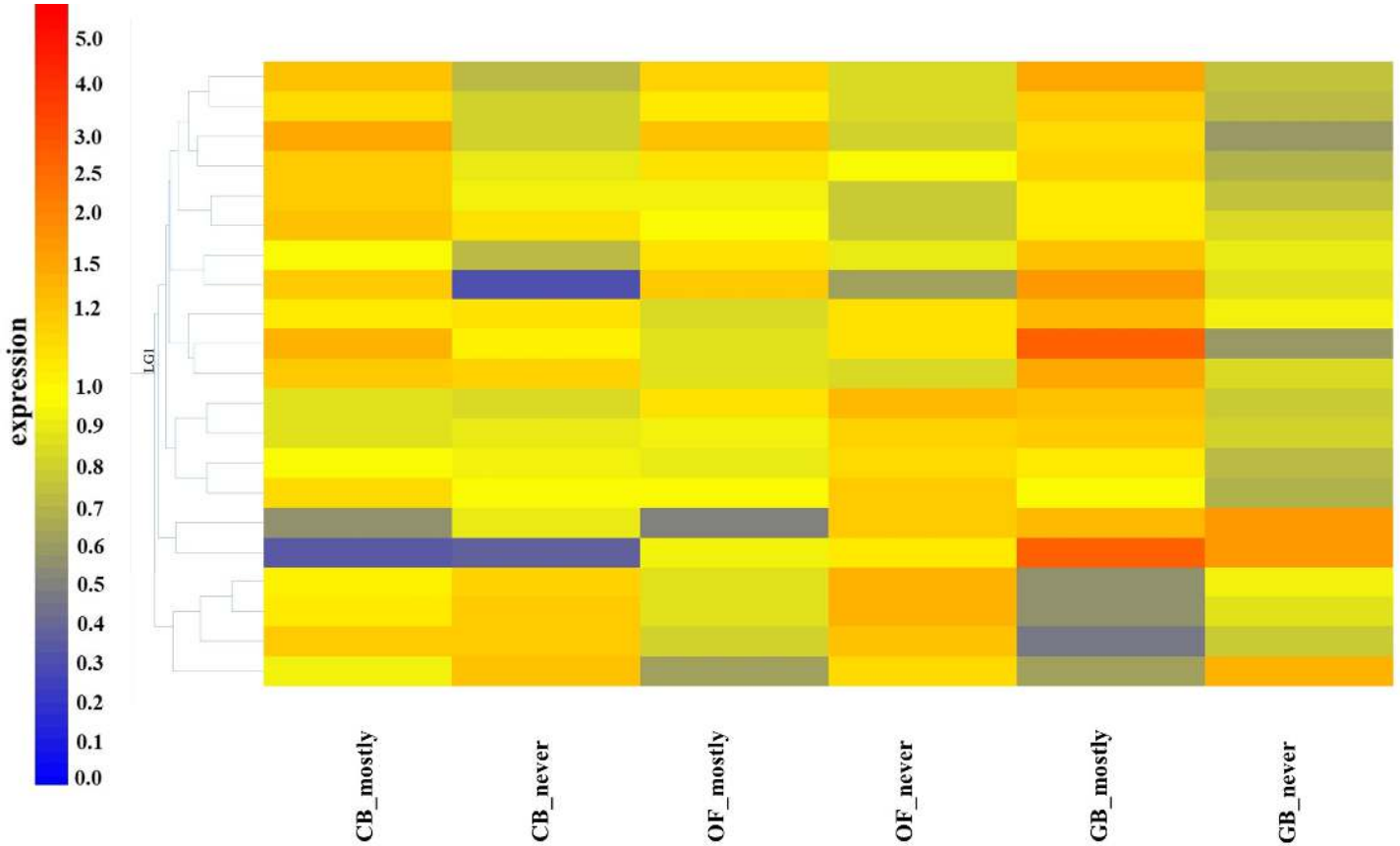


Figure 4.4: Tree cluster heatmap of the microarray probes mapped on the 'crumbly' QTL on LG1.

The twenty-one probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') for the three different development stages (closed bud (**CB**), open flower (**OF**) and green berry (**GB**) with differences being significative, with respect to stage*phenotype interaction, at 99.9% confidence levels. All the probes were mapped inside the 'crumbly' QTL on linkage group 1 (LG1). High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented.

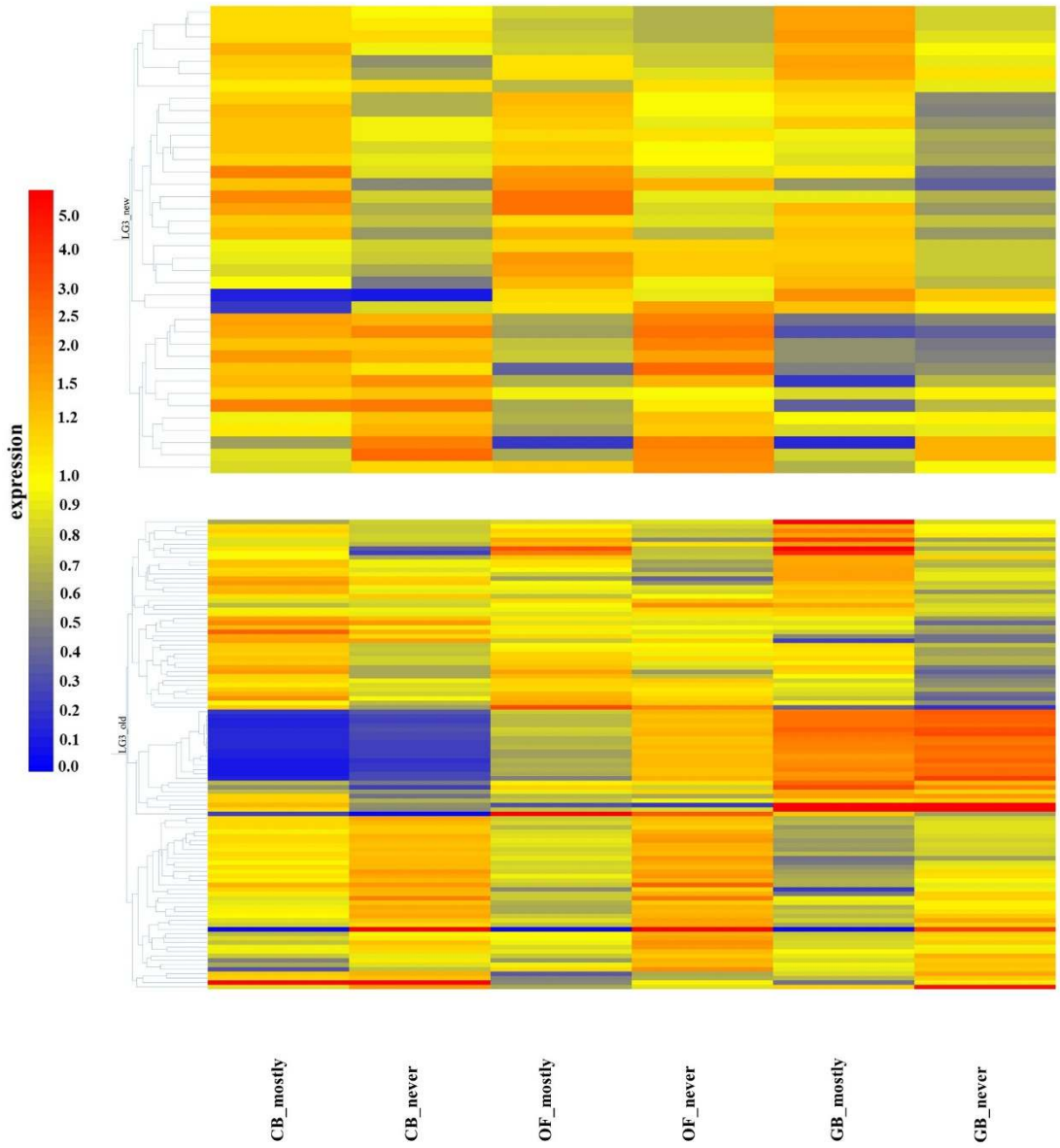


Figure 4.5: Tree cluster heatmaps of the microarray probes mapped on the two 'crumbly' QTLs on LG3.

Top figure, the thirty-eight probes mapped inside the 'crumbly' QTL, identified during this work. Bottom figure, the one hundred six probes mapped inside the original 'crumbly' QTL (Graham et al. 2015). All the probes were mapped on linkage group 3 (LG3), they all were differently expressed, with respect to stage*phenotype interaction, between the two phenotypes (i.e. mostly and never 'crumbly') for the three different development stages (closed bud (**CB**), open flower (**OF**) and green berry (**GB**) with differences being significative at 99.9% confidence levels. High expression levels are indicated in red colour while low expression level in blue colour as per scale bar presented.

4.3.4 Gene ontology annotation of differently expressed probes mapped in the three crumbly QTLs

All the 165 probes within the QTLs, with expression levels significantly differentially expressed in the microarray experiment, were analysed individually through the GO bulk TAIR database to look for gene annotation as previously undertaken for the selected probes in chapter three. Here the analysis of genes within the QTL was conducted as in chapter three, and specifically focused on annotation terms related to flower development, hormones, pollen and transport.

The analysis of the 165 *A. thaliana* ortholog genes of those *R. idaeus'* ones matched by the 'crumbly' microarray probes, mapped inside the 'crumbly' QTLs, and significantly differently expressed with respect to stage*phenotype interaction effect, allowed the selection of only seventeen predicted genes. These genes have ontology annotation related to flower development, hormones and pollen, they were selected and listed in Tables 4.4 and 4.5. Thirteen of the seventeen selected probes matched predicted *Arabidopsis thaliana* genes whose gene ontology terms were specifically related to hormones with reference to their metabolism, modulation and response. Cytokinins, auxins and jasmonates were the most represented, having four different annotations, they were followed by salicylates with three annotations, ABA with two and gibberellin and steroids with only one. Together with the hormones, two other classes of annotations were considered, flower development and pollen. Regarding the probes matching predicted genes with gene ontology terms related to pollen, these were two and the specific annotation were 'recognition of pollen and 'pollen development'. Moreover, three probes matched as many predicted genes whose annotations terms, related to flower development, were reported as for all the other significant probes on Tables 4.4 and 4.5 while on Tables 4.6, 4.7 and 4.8 the ANOVA table of means for each of the selected probes was reported. In Figure 4.6 a tree cluster heatmap related to these sixteen selected microarray probes was reported.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	*LG
CUST_10154_PI426541283	cytokinin metabolic process	GO:0009690	AT3G63440.1	3_new
	cytokinin dehydrogenase activity	GO:0019139		
CUST_13398_PI426541283	abscisic acid metabolic process	GO:0009687	AT2G04240.1	3
	response to gibberellin	GO:0009739		
CUST_18787_PI426541283	response to auxin	GO:0009733	AT2G04160.1	3
CUST_20489_PI426541283	response to cytokinin	GO:0009735	AT1G11910.1	3
CUST_22099_PI426541283	steroid binding	GO:0005496	AT5G52240.1	1
CUST_27324_PI426541283	cytokinin dehydrogenase activity	GO:0019139	AT3G63440.1	3_new
	cytokinin metabolic process	GO:0009690		
CUST_28007_PI426541283	[a]MeIAA esterase activity	GO:0080030	AT3G29770.1	3
	[b]MeSA esterase activity	GO:0080031		
	[c]MeJA esterase activity	GO:0080032		
	jasmonic acid metabolic process	GO:0009694		
	salicylic acid metabolic process	GO:0009696		
CUST_33454_PI426541283	auxin activated signalling pathway	GO:0009734	AT1G05180.1	3
	auxin homeostasis	GO:0010252		
	response to cytokinin	GO:0009735		

[a]methyl indole-3-acetate – [b]methyl salicylate – [c]methyl jasmonate
[*]linkage group

Table 4.4: : List of microarray probes mapped inside the three 'crumbly' QTLs on linkage group 1 and 3.

List of the probes matching Arabidopsis thaliana predicted genes with gene ontology (GO) annotation related to hormones or pollen or flower development. LG3_new indicates the 'crumbly' QTL identified during this work on linkage group 3 while LG1 and LG3 indicate the original 'crumbly' QTLs identified by Graham et al. (2015) on linkage groups 1 and 3 respectively.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID	*LG
CUST_37835_PI426541283	response to auxin	GO:0009733	AT1G30330.2	3
	auxin activated signalling pathway	GO:0009734		
CUST_38171_PI426541283	response to abscisic acid	GO:0009737	AT3G57540.1	3_new
CUST_38657_PI426541283	response to jasmonic acid	GO:0009753	AT2G46410.1	3
	response to salicylic acid	GO:0009751		
CUST_44619_PI426541283	response to cyclopentenone (OPDA)	GO:0010583	AT3G12110.1	3_new
CUST_54460_PI426541283	cytokinin biosynthetic process	GO:0009691	AT1G62360.1	3
	carpel development	GO:0048440		
	floral meristem determinacy	GO:0010582		
CUST_12159_PI426541283	recognition of pollen	GO:0048544	AT2G19130.1	3
CUST_16314_PI426541283	plant ovule development	GO:0048481	AT3G55400.1	3_new
CUST_24407_PI426541283	pollen development	GO:0009555	AT5G12210.1	3
CUST_35866_PI426541283	embryo development ending in seed dormancy	GO:0009793	AT4G28210.1	3

*linkage group

Table 4.5: List of microarray probes mapped inside the three 'crumbly' QTLs on linkage group 1 and 3.

List of the probes matching Arabidopsis thaliana predicted genes with gene ontology (GO) annotation related to hormones or pollen or flower development. LG3_new indicates the 'crumbly' QTL identified during this work on linkage group 3 while LG1 and LG3 indicate the original 'crumbly' QTLs identified by Graham et al. (2015) on linkage groups 1 and 3 respectively.

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009690 [b](cyt. metab. process) GO:0019139 [c](cyt. dehyd. activity)	AT3G63440.1	mostly	0.055	0.03*	0.205	0.0485	CUST_10154_PI426541283[k]
		never	-0.185	-0.044*	0.027		
GO:0009687 (ABA metabolic process) GO:0009739 (response to gibberellins)	AT2G04240.1	mostly	-0.503	0.048	-0.037*	0.0705	CUST_13398_PI426541283[j]
		never	-0.11	0.272	0.091*		
GO:0009733 (response to auxin)	AT2G04160.1	mostly	-0.007*	0.107*	0.197	0.0632	CUST_18787_PI426541283[j]
		never	-0.177*	-0.093*	0.055		
GO:0009735 (response to cytokinin)	AT1G11910.1	mostly	0.092	0.076*	0.006	0.0373	CUST_20489_PI426541283[j]
		never	-0.077	-0.032*	-0.163		
GO:0005496 (steroid binding)	AT5G52240.1	mostly	0.082	0.08*	0.25*	0.1526	CUST_22099_PI426541283[h]
		never	-0.497	-0.187*	-0.044*		
GO:0009690 [b](cyt. metab. process) GO:0019139 [c](cyt. dehyd. activity)	AT3G63440.1	mostly	0.08	0.02*	0.215	0.0407	CUST_27324_PI426541283[k]
		never	-0.233	-0.099*	-0.029		
GO:0080030 (MeIAA esterase activity) GO:0080031 (MeSA esterase activity) GO:0080032 (MeJA esterase activity) GO:0009694 (JA metabolic process) GO:0009696 (SA metabolic process)	AT3G29770.1	mostly	0.059*	0.135*	-0.001	0.0907	CUST_28007_PI426541283[j]
		never	-0.176*	-0.123*	-0.33		

[a]2 degrees of freedom and 4 replicates
*difference statistically not significant
[h]QTL on linkage group 1 – [J]QTL on linkage group 3 previously identified by Graham et al., (2015) – [k]QTL on linkage group 3 recently identified during this work
[b]cytokinin metabolic process – [c]cytokinin dehydrogenase activity

Table 4.6: ANOVA table of means (stage*phenotype interaction) of microarray probes mapped inside the three 'crumbly' QTLs on linkage group 1 and 3.

List of the probes matching Arabidopsis thaliana predicted genes with gene ontology (GO) annotation related to hormones or pollen or flower development. The probes mapped inside the three 'crumbly QTLs, one on linkage group 1 and two on linkage group 3. ANOVA table of means for the interaction phenotype (i.e. mostly and never 'crumbly') per stage (i.e. closed bud, open flower and green berry).

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009734 [b](auxin act. sig. path.) GO:0019139 (auxin homeostasis) GO:0009735 (response to cytokinin)	AT1G05180.1	mostly	0.097	0.07*	0.042	0.0441	CUST_33454_PI426541283[j]
		never	-0.077	0.037*	-0.152		
GO:0009733 (response to auxin) GO:0009734 [b](auxin act. sig. path.)	AT1G31330.2	mostly	0.021*	-0.037	-0.359*	0.0646	CUST_37835_PI426541283[j]
		never	0.145*	0.241	-0.201*		
GO:0009737 (response to abscisic acid)	AT3G57540.1	mostly	-0.856	0.044*	0.297	0.0392	CUST_38171_PI426541283[k]
		never	-1.051	-0.03*	0.092		
GO:0009753 (response to JA) GO:0009751 (response to SA)	AT2G46410.1	mostly	0.063*	0.056*	0.334	0.0598	CUST_38657_PI426541283[j]
		never	-0.104*	-0.114*	-0.008		
GO:0005496 [c](response to cyclop.)	AT3G12110.1	mostly	-0.2	-0.669	-0.765	0.1335	CUST_44619_PI426541283[h]
		never	0.34	0.341	0.159		

[a]2 degrees of freedom and 4 replicates
*difference statistically not significant
[h]QTL on linkage group 1 – [J]QTL on linkage group 3 previously identified by Graham et al., (2015) – [k]QTL on linkage group 3 recently identified during this work
[b]auxin activated signalling pathway – [c]response to cyclopentenone (OPDA)

Table 4.7: ANOVA table of means (stage*phenotype interaction) of microarray probes mapped inside the three 'crumbly' QTLs on linkage group 1 and 3.

List of the probes matching Arabidopsis thaliana predicted genes with gene ontology (GO) annotation related to hormones or pollen or flower development. The probes mapped inside the three 'crumbly QTLs, one on linkage group 1 and two on linkage group 3. ANOVA table of means for the interaction phenotype (i.e. mostly and never 'crumbly') per stage (i.e. closed bud, open flower and green berry).

Gene Ontology (GO) term	*A. thaliana* gene ID	Crumbly	stage expression levels			[a]S.E. crumbly/stage	Probe microarray
			closed	open	green		
GO:0009691 [b](cytokinin biosynt. proc.) GO:0048440 (carpel development) GO:0010582 [c](floral meristem determ.)	AT1G62360.1	mostly	-0.14	0.007*	0.435	0.0454	CUST_54460_PI426541283[j]
		never	-0.324	0.021*	0.133		
GO:0048544 (recognition of pollen)	AT2G19130.1	mostly	0.183*	0.104*	0.073	0.1093	CUST_12159_PI426541283[j]
		never	-0.167*	-0.06*	-0.316		
GO:0048481 (plant ovule development)	AT3G55400.1	mostly	0.057*	0.067*	-0.056	0.0373	CUST_16314_PI426541283[k]
		never	-0.036*	0.001*	-0.173		
GO:0009555 (pollen development)	AT5G12210.1	mostly	-0.041*	-0.148	-0.032*	0.0487	CUST_24407_PI426541283[j]
		never	0.105*	0.077	0.005*		
GO:0009793 [d](embryo dev. ending in)	AT4G28210.1	mostly	0.023	-0.028	-0.172	0.0582	CUST_35886_PI426541283[j]
		never	0.187	0.24	0.026		

[a]2 degrees of freedom and 4 replicates
*difference statistically not significant
[h]QTL on linkage group 1 – [J]QTL on linkage group 3 previously identified by Graham et al., (2015) – [k]QTL on linkage group 3 recently identified during this work
[b]cytokinin biosynthetic process – [c]floral meristem determinacy – [d]embryo development ending in seed dormancy

Table 4.8: ANOVA table of means (stage*phenotype interaction) of microarray probes mapped inside the three 'crumbly' QTLs on linkage group 1 and 3.

List of the probes matching Arabidopsis thaliana predicted genes with gene ontology (GO) annotation related to hormones or pollen or flower development. The probes mapped inside the three 'crumbly QTLs, one on linkage group 1 and two on linkage group 3. ANOVA table of means for the interaction phenotype (i.e. mostly and never 'crumbly') per stage (i.e. closed bud, open flower and green berry).

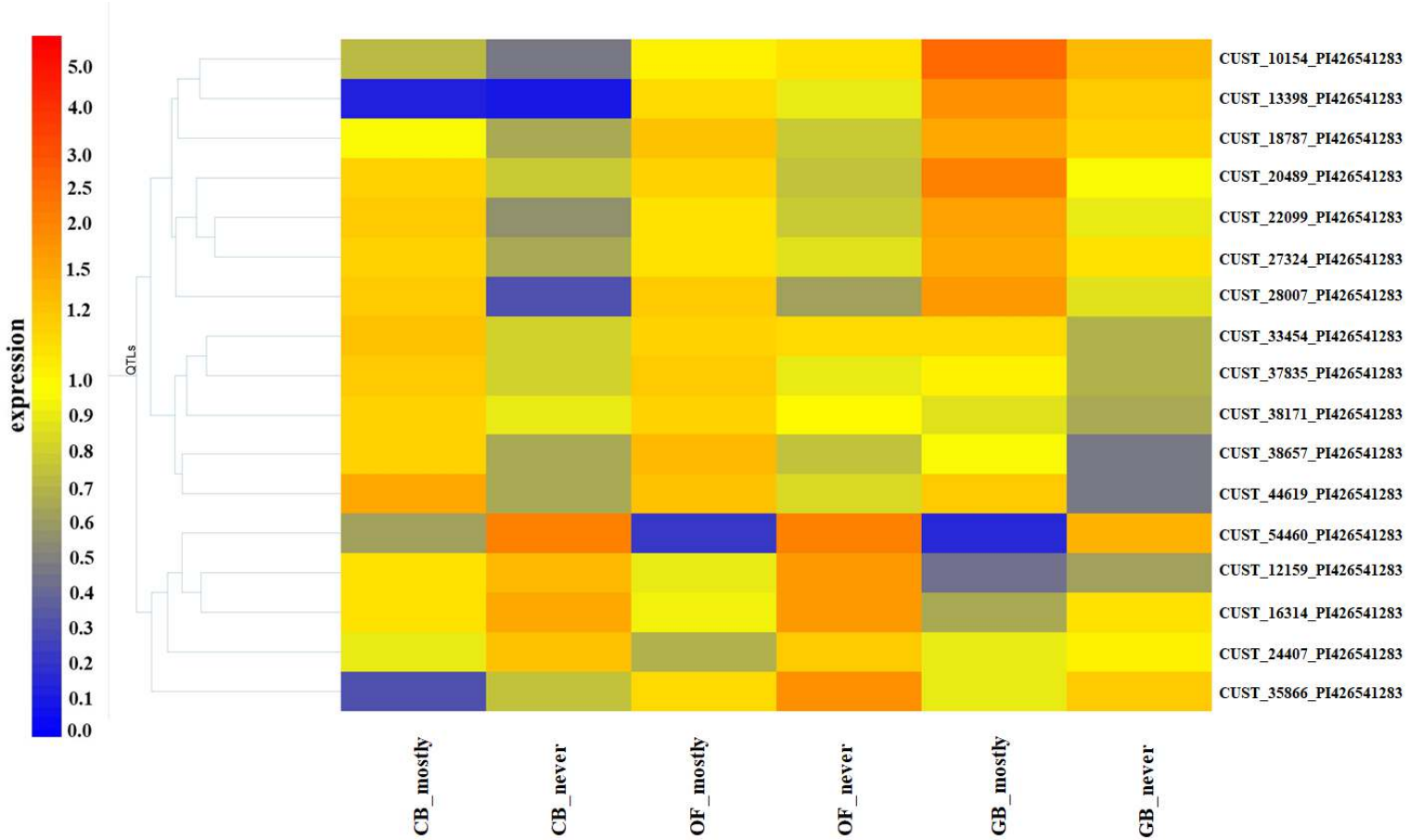


Figure 4.6: Tree cluster heatmaps of sixteen microarray probes mapped inside the three 'crumbly' QTLs.

The seventeen probes were specifically selected through gene ontology annotation analysis because they have specific ontology terms related to flower development (3 probes), pollen (2 probes) and hormones (12 probes).

4.3.5 Relationship between 'crumbly' and fruit ripening

In their work, Graham et al., (2015) found correlations between the 'crumbly' scores and those related to fruit ripening. The largest correlations were found with the time to reach fruit set and that to reach green fruit. These correlations were positive, meaning that the proportion of 'crumbly' fruit increased with the time plants took to reach these two development stages. The longer the fruits take to get to fruit set and green fruit, the more likely they tend to be 'crumbly' (Graham et al., 2015). Ripening is associated with many markers and many different QTLs were found (Graham et al., 2009). The ripening score consisted of repeated observations carried out on a finite number of occasions. The scores on different dates were then combined using principal coordinates (PCOs) to extract the principal sources of variation in ripening across the mapping population (Graham et al., 2009). The PCOs, being uncorrelated to each other, allowed an easy allocation of the QTLs but the interpretation of the QTLs in respect to the original ripening stage was very hard. It was very difficult to interpret the principal coordinates in terms of the effect of marker genotypes on the ripening process. For such reason a second approach was adopted by interpolating the time to reach each stage of ripening. This strategy allowed the natural interpretation of the QTL effect by simply detecting the presence or absence of a marker in the number of days to reach a certain ripening stage (Graham et al., 2009).

The same ripening scores were re-analysed with the new GbS linkage map and 12 of these QTLs were confirmed (Hackett et al. 2018). The overlapping regions between fruit ripening and 'crumbly' QTLs were analysed to look for association and of particular interest seemed to be both the two 'crumbly' QTLs identified on LG3. The QTL, identified in this work, having **Rub2a 1** as most representative marker and located at 62.3 cM (centimorgan) and the previous QTL with **s182_p91185_R6** as most representative marker and located at 106.4 cM also impact ripening.

The rule of thumb to define the size of each QTL was the one-LOD support interval (Hackett, 2002) which allowed the identification of the confidence intervals for each QTL by selecting a region on both sides of the QTL peak that corresponded to a decrease of 1 LOD score. The estimated supporting intervals for each 'crumbly' QTLs were: 1-12 cM for the QTL on LG1 and 104-113 cM for the other QTL identified by Graham et al. (2015). For the 'crumbly' QTL, identified in this work by re-analysing the 'crumbly' phenotypic scores (Graham et al., 2015) with the new Genotyping by Sequencing (GbS) map (Hackett et al., 2018) the QTL support interval was 53-64 cM.

The 'crumbly' QTL identified in this study overlapped two ripening QTLs, representing the open flower stage of ripening (62 cM) and the percentage of open flower (66 cM) at a defined stage. The 'crumbly' QTL previously identified, overlapped four QTLs, three of them referring to trials in the open field (i.e. fruit set, green fruit and PCO2), respectively at 102, 105 and 94 cM, and the fourth fruit set again but this time for polytunnel trials at 102 cM (Hackett et al., 2018).

Principal coordinate 2 (PCO2) was highly positive correlated with late June and early July ripening scores and had a negative correlation with the ripening scores in May meaning that positive scores for this principal coordinate indicated a slow development in May followed by a fast ripening in late June and early July (Graham et al., 2009). All the QTLs mapped on LG3 were mainly caused by alleles from the Latham parent, except for the % of open flower in the polytunnel trails where the effect of both parents on the phenotype was significative (Hackett et al., 2018). In Tables 4.9 and 4.10, the scaffolds comprised in the overlapping regions of the fruit ripening QTLs and the two 'crumbly' QTLs, respectively, the one recently identified during this work and that previously identified by Graham et al., (2015) were reported.

LG3 identified during this work	**Trait**	**'crumbly'**	**fruit ripening**	
	trials	**field**	**polytunnel**	
	QTL	**'crumbly'**	**open flower**	**%open flower**
	***cM**	**62.3**	**62**	**66**
		scaffolds		
		367		**367**
		499	**499**	

*centimorgan

Table 4.9: : Scaffolds comprised in the overlapping 'crumbly' and fruit ripening QTLs.

List of shared scaffolds within the 'crumbly' QTL recently identified during this work and those of fruit ripening identified by Hackett et al., (2018) all on linkage group 3. The two different ripening QTLs (i.e. open flower and % of open flower) refer to the time to reach respectively open flower and the percentage of flower opened at the middle of June (i.e. the 13th of June).

	Trait	'crumbly'	fruit ripening			
	trials	field	field			[l]poly.
	QTL	'crumbly'	[a]PCO2	[§]g. fruit	fruit set	fruit set
	[*]cM	106.4	94	102	105	102
LG3 previously identified by Graham et al., (2015)		scaffold				
		4		4	4	
		43		43		43
		52			52	
		65		65	65	
		76	76			
		182		182		
		294			294	
		664	664			
		734		734	734	
		835	835			
		858		858	858	

[*]centimorgan
[l]polytunnel
[a]principal coordinate 2
[§]green fruit

Table 4.10: Scaffolds comprised in the overlapping 'crumbly' and fruit ripening QTLs.

List of shared scaffolds within the 'crumbly' QTL previously identified by Graham et al., (2015) and those of fruit ripening identified by Hackett et al., (2018) all on linkage group 3. The two different ripening QTLs (i.e. fruit set and green) refer to the time to reach respectively fruit set and green fruit while PCO2 indicates the second source of variation extracted from the overall summary of all the ripening scores (i.e. repeated observations carried out on a finite number of occasions) combined together using principal coordinates analysis (Graham et al., 2009).

The 'crumbly' microarray experiment described in chapter three of this work allowed the identification of genes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') for a sample of 14 individuals of the Glen Moy x Latham mapping population (see chapter 3 for further details) across the whole Glen Moy genome. In this chapter we identified microarray probes within scaffolds in QTLs regions identified from the whole mapping population that were differentially expressed. Both approaches were

used to ensure as many significant differentially expressed genes as possible were identified including those of minor effects.

Five microarray probes from the 'crumbly' experiment described in chapter 3 matched *R. idaeus* genes mapped in overlapping regions of fruit ripening and 'crumbly' QTLs on linkage group 3 as reported on Table 4.11. In Table 4.12, the predicted mean form the ANOVA stage*phenotype interaction were reported to show the significant (0.001 confidence levels) differences in their expression levels.

The probe CUST_54460_PI426541283 was differently expressed with the difference being statistically significative in the closed bud and green berry stages but not in the open flower stage (see Table 4.12). The probe, in the mostly 'crumbly' phenotype, was upregulated at both closed bud and green berry stage and the corresponding ortholog *Arabidopsis thaliana* gene (AT1G62360.1) was mapped inside the old 'crumbly' QTL as well as with the fruit ripening QTLs, fruit set and green fruit. Two important gene ontology annotations were found for this gene, GO:0048440 (carpel development) and GO:0009691 (cytokinin biosynthetic process).

The second interesting 'crumbly' microarray probe CUST_35866_PI426541283, matched a gene mapped inside the scaffold65 that was located within the old 'crumbly' QTL and again co-located with two fruit ripening QTLs, fruit set and green fruit. The probe was significatively differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with 99.9% confidence levels. The probe was downregulated in the mostly 'crumbly' phenotype in all the three stages tested (i.e. closed bud, open flower and green berry). The probe matched the ortholog *Arabidopsis thaliana* gene (AT4G28210.1) that had two interesting gene ontology annotations, GO:0009793 (embryo development ending in seed dormancy) and GO:0009409 (response to cold); the full details for this probe were reported in Table 4.11.

					trait				
					‘crumbly’	fruit ripening			
Microarray probe	***A. thaliana* gene ID**	**[a]cM**	**[b]LG**	**scaffold**	**QTLs (field trials**				**[d]G.O. terms**
					‘crumbly’	**[c]PCO2**	**fruit set**	**green fruit**	
CUST_54460_PI426541283	AT1G62360.1	99.4	3	4	**X**		**X**	**X**	[1]GO:0048440 [2]GO:0009691
CUST_35866_PI426541283	AT4G28210.1	103.3	3	65	**X**		**X**	**X**	[3]GO:0009793 [4]GO:0009409
CUST_37835_PI426541283	AT1G30330.2	84.3	3	664	**X**	**X**			[5]GO:0009734; [6]GO:0009733; [7]GO:0009808
CUST_24407_PI426541283	AT5G12210.1	102	3	734	**X**		**X**	**X**	GO:0009555 (pollen development)
CUST_12159_PI426541283	AT2G19130.1	105	3	858	**X**		**X**		GO:0048544 (recognition of pollen)

[1]carpel development – [2]cytokinin biosynthesis – [3]embryo development ending in seed dormancy – [4]response to cold – [5]auxin activated signalling pathway – [6]response to auxin – [7]flower development – [a]centimorgan – [b]linkage group – [c]principal coordinate 2 – [d]gene ontology

Table 4.11: List of ‘crumbly’ microarray probes mapped in the overlapping regions of the ‘crumbly’ and fruit ripening QTLs on linkage group 3.

For each probe was reported the matched *A. thaliana* ortholog gene and the scaffold were the *Rubus idaeus* equivalent gene is located. The last column reports the gene ontology terms related to the orthologs Arabidopsis thaliana genes matching the probes being analysed in this work.

Microarray probe	A. thaliana gene ID	'crumbly'	stage			[a]S.E. crumbly/stage
			closed	open	green	
CUST_54460_PI426541283	AT1G62360.1	mostly	-0,14	0.007[b]	0.435	0.0454
		never	-0.324	0.021[b]	0.133	
CUST_35866_PI426541283	AT4G28210.1	mostly	0.023	-0.028	-0.172	0.0582
		never	0.187	0.24	0.026	
CUST_37835_PI426541283	AT1G30330.2	mostly	0.021[b]	-0.037	-0.359[b]	0.0646
		never	0.145[b]	0.241	-0.201[b]	
CUST_24407_PI426541283	AT5G12210.1	mostly	-0.041[b]	-0.148	-0.032[b]	0.0487
		never	0.105[b]	0.077	0.005[b]	
CUST_12159_PI426541283	AT2G19130.1	mostly	0.183[b]	0.104[b]	0.073	0.1093
		never	-0.167[b]	-0.06[b]	-0.316	

[*]ANOVA with 2 degrees of freedom and 4 replicates
[b]difference being non statistically significant at 99.9% confidence levels

Table 4.12: List of 'crumbly' microarray probes mapped in the overlapping regions of the 'crumbly' and fruit ripening QTLs on linkage group 3.

ANOVA table of the means for the interaction phenotype (i.e. mostly and never 'crumbly) x stage (i.e. closed bud open flower and green berry) for the 'crumbly' microarray probes matching scaffolds mapped within the old 'crumbly' QTL and the fruit ripening QTLs related to the field trials. on LG3.

The 'crumbly' QTL, previously identified by Graham et al., (2015) overlapped with the PCO2 fruit ripening QTL for the field trials with scaffold664 underlying these. The probe CUST_37835_PI426541283 matched scaffold664 and the *Arabidopsis thaliana* ortholog gene AT1G30330.2. The gene was associated with three interesting gene ontology terms, GO:0009734 (auxin activated signalling pathway), GO:0009733 (response to auxin); GO:0009808 (flower development). The probe, in the 'crumbly' microarray experiment, was differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with the difference being significant at 99.9% confidence level only for open; the probe was downregulated in the mostly 'crumbly' samples (see Table 4.12 for full details).

The last two probes, CUST_24407_PI426541283 and CUST_12159_PI426541283 were both mapped inside the previously identified 'crumbly' QTL (Graham et al., 2015) as well co-locating with two fruit ripening QTLs, fruit set and green fruit (field trials). While the first probe matched a gene inside a scaffold with two ripening QTLs (i.e. fruit set and green fruit), the second probe co-located with only one QTL for fruit set (see Tables 4.9 and 4.10 for full details). The overlapping genome regions between these QTLs were respectively, scaffold734 for the first probe and scaffold858 for the second one. The probes matched the *Arabidopsis thaliana* ortholog genes, respectively AT5G12210.1 and AT2G19130.1. The two genes both presented gene ontology annotations related to pollen and more in details, for the first gene GO:0009555 (pollen development) and GO:0048544 (recognition of pollen) for the second gene. The microarray probes were again differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significative only at open stage for the first probe while only at green berry for the second probe (see Table 4.12).

4.3.6 Steps towards the selection of 'crumbly' genetic markers for breeding assisted and diagnostic purposes

In total, eight potential 'crumbly' fruit loci were selected for investigation for marker development. These loci were selected simply for being the most representative marker or at least its closest neighbour within the QTL. Five of these loci were mapped as Single Nucleotide polymorphisms (SNPs), while the other three, previously characterised, were Simple Sequence Repeats (SSRs). The eight gene markers selected were of two different kinds; the so called gene 'tags', located in non-coding regions of the genome, and the

target genes, on the other hand, located in coding ones (Collard et al., 2005). **MOY_34151, MOY_35728** and **MOY_36258** were the only genes targets (see Tables 4.13 to 4.16 for further details), all the rest were genes 'tag'. These three genes were located on LG1 (**MOY_34151**) while the other two were both on LG3 one in the QTL identified here (**MOY_35728**) and the other in the original **'**crumbly' QTL (**MOY_36258**).

Two SNPs markers, **s3407_p12510_R23** and **s182_p91185_R6,** the two new most significant 'crumbly' markers for, respectively, the QTL on LG1 and the QTL on LG3, were identified during this work by re-analysing the 'crumbly' score (Graham et al., 2015) with the new GbS linkage map (Hackett et al., 2018).

The last three genes were all SSRs, **Rub2a1** was the most significant marker for the 'crumbly' QTL here identified during this work while **ERubLR_SQ05.3_D11AOC** and **Rub256e** were the original most significant 'crumbly' markers, respectively for the QTL on LG3 and the one on LG1; identified by Graham et al., (2015) before the development of the new GbS linkage map (Hackett et al., 2018). These two SSR markers were also considered for the Genome Wide Association Study (GWAS) because they were inside their corresponding 'crumbly' QTL and they were even tightly linked to the new most significative markers. Therefore, the principle followed was, the more markers selected and tested the more chances to find at least one 'crumbly' marker being representative of a wide population of individuals showing 'crumbly' symptoms.

The design of the primers for the marker (**s3407_p12510_R23)** failed due to missing sequences in the **scaffold3407** where it is mapped. A new SNP marker (**s353_p21288_R19**), tightly linked with (**s3407_p12510_R23)** was then used.

The primers for all the eight markers were designed as described in section 4.2.3. They were used for a validation pool whose aim was to identify a 'crumbly' marker either always or never linked to the 'crumbly' population analysed. Ideally the marker/s should be shared by all the genotypes labelled as 'crumbly' (45 genotypes in this case study) while being absent in all those that never displayed 'crumbly' symptoms (18 genotypes in this case study) or vice versa (see Table A.4.1 for full list of genotypes).

The identification of the five SNPs markers, the two gene tags and the three gene targets, was performed by sequencing the amplified regions containing the markers across the genotypes. The procedure accomplished was described in section 4.2.4 by using the primers designed as described in section 4.2.3. The results of the first sequencing analysis,

performed only with the forward primers, did not identify any SNP marker that was always (or never) found in all the 45 genotypes, having 'crumbly' phenotype, out of the 63 genotypes (see full list of genotypes in appendix Table A.4.1) tested. Fourteen, out of the sixty-three, genotypes selected for the GWAS did not produce any amplification fragments; suggesting that their genome sequence differed in the primer design regions (see Table 4.17 for further details) therefore these were not informative in these genotypes.

The amplified region sequenced, with the procedure described in section 4.2.4, for the five different markers varied slightly between the genotypes. The contig, the set of overlapping DNA sequence for the different genotypes varied in length indicated in base pair (bp): **MOY_34151** (235 bp), **MOY_36258** (186 bp), **MOY_35728** (285 bp), **s182_p91185_R6** (487 bp) and **s353_p21288_R19** (355 bp). The number of SNPs per sequence identified varied per locus. The contig for **MOY_34151** had four different SNP positions 139 bp, 146 bp, 158 bp and 183 bp. The contig for **MOY_36258** had five different SNP positions but none segregating in the selected population. The contig for **MOY_35728** carried three different SNP positions (i.e. 75 bp, 174 bp and 242 bp). The contig for the locus **s182_p91185_R6** had eleven different SNP positions (i.e. 110 bp, 129 bp, 139 bp, 147 bp, 148 bp 170 bp, 206 bp, 236 bp, 376 bp, 441 bp and 444 bp). The contig of the last marker, **s353_p21288_R19**, carried only three different SNP positions (i.e. 180 bp, 196 bp and 303 bp).

Probe microarray	*A. thaliana* gene ID	[a]LG	Crumbly	stage expression levels			[b]S.E. crumbly/stage	Glen Moy gene ID
				closed	open	green		
CUST_25407_PI426541283	AT5G02790.1	1	mostly	0.117	-0.001*	0.0152	0.03018	MOY_34151
			never	0.0266	-0.094*	-0.068		
CUST_35082_PI426541283	AT5G02790.1	1	mostly	0.0791*	-0.014*	0.0114	0.0336	MOY_34151
			never	-0.02*	-0.105*	-0.111		
CUST_52360_PI426541283	AT5G02790.1	1	mostly	0.037	0.018*	0.073	0.035	MOY_34151
			never	-0.08	-0.06*	-0.132		
CUST_6744_PI426541283	AT5G02790.1	1	mostly	0.101	0.07	0.183	0.0376	MOY_34151
			never	-0.143	-0.068	-0.123		

[a]linkage group 'crumbly' QTL identified by Graham *et al*., (2015) on linkage group 1
[b]2 degrees of freedom and 4 replicates
*difference statistically not significant

Table 4.13: List of 'crumbly' microarray probes matching the selecting marker (genes target) highly associated with the most significant markers for the 'crumbly' QTL identified on linkage group 1 previously by Graham *et al*., (2015).

List of the 'crumbly' microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and mapped close to the most significant genotype marker of the 'crumbly' QTL on linkage group 1.

Probe microarray	*A. thaliana* gene ID	[a]LG	Crumbly	stage expression levels			[b]S.E. crumbly/stage	Glen Moy gene ID
				closed	open	green		
CUST_28007_PI426541283	AT3G29770.1	3	mostly	0.059*	0.135*	-0.001	0.0907	[1]MOY_36258
			never	-0.176*	-0.123*	-0.33		
CUST_52_PI426541283	AT3G29770.1	3	mostly	-0.13	0.001	-0.093	0.041	[1]MOY_36258
			never	-0.008	0.172	0.053		
CUST_26373_PI426541283	AT2G21540.1	3	mostly	0.241*	-0.103	-0.245*	0.0649	[2]MOY_35728
			never	0.159*	0.218	-0.302*		
CUST_36699_PI426541283	AT2G21540.1	3	mostly	0.184*	-0.203	-0.495*	0.1126	[2]MOY_35728
			never	0.328*	0.419	-0.409*		
CUST_48963_PI426541283	AT2G21540.1	3	mostly	0.113*	-0.121	-0.243*	0.07	[2]MOY_35728
			never	0.119*	0.339	-0.311*		
CUST_49921_PI426541283	AT2G21540.1	3	mostly	0.228*	-0.185	-0.348*	0.0936	[2]MOY_35728
			never	0.157*	0.334	-0.269*		

[a]linkage group
[b]2 degrees of freedom and 4 replicates
*difference statistically not significant
[1]'crumbly' QTL recently identified during this work on linkage group 3
[2]'crumbly' QTL previously identified by Graham *et al*., (2015)

Table 4.14: List of 'crumbly' microarray probes matching the two selecting markers (genes target) highly associated with the most significant markers for the two 'crumbly' QTLs on linkage group 3, the one identified previously by Graham *et al*., (2015) and that recently identified in this work.

List of the 'crumbly' microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and mapped close to the most significant genotype marker of the two 'crumbly' QTLs on linkage group 3.

Probe microarray	scaffold	[a]cM	[b]LG	probe position	Glen Moy gene ID	position gene along scaffold	[c]SNPs
CUST_25407_PI426541283	462	6.1	1	54797 - 54855	MOY_34151	53465 - 56776	53484 – 53624 – 53981 – 54196 – 54369 – 54680 – 55085 – 55088 – 55380 – 55473 – 55648 – 55856 – 56244 – 56342 – 56351 – 56381 – 56618 - 56676
CUST_35082_PI426541283	462	6.1	1	54805 - 54863	MOY_34151		
CUST_52360_PI426541283	462	6.1	1	55323 - 55382	MOY_34151		
CUST_6744_PI426541283	462	6.1	1	56327 - 56386	MOY_34151		

[a]centimorgan
[b]linkage group
[c]single nucleotide polymorphism

Table 4.15: List of 'crumbly' microarray probes matching the selecting marker (genes target) highly associated with the most significant markers for the 'crumbly' QTL identified on linkage group 1 previously by Graham *et al*., (2015).

List of the 'crumbly' microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and mapped close to the most significant genotype marker of the 'crumbly' QTL on linkage group 1. For each probe was reported the scaffold to which the probe matched, the position both in cM and in base pair along the scaffold. Moreover, the position of the Glen Moy marker (gene target) along the scaffold and all its SNPs with their position in base pair were reported too.

Probe microarray	scaffold	[a]cM	[b]LG	probe position	Glen Moy gene ID	position gene along scaffold	[c]SNPs
CUST_28007_PI426541283	509	110.2	3	29834 – 29893	MOY_36258	28806 - 32917	29173 – 29197 – 29284 – 29676 – 29752 – 29868 – 30209 – 30315 – 30432 – 30541 – 30565 - 30736
CUST_52_PI426541283	509	110.2	3	32791 - 32850	MOY_36258		
CUST_26373_PI426541283	499	61.2	3	73838 – 73897	MOY_35728	67983 - 74349	68630 – 70107 – 72215 – 72534 – 72611 – 72760 – 72773 – 73638 – 73733
CUST_36699_PI426541283	499	61.2	3	70083 – 70142	MOY_35728		
CUST_48963_PI426541283	499	61.2	3	72569 – 72628	MOY_35728		
CUST_49921_PI426541283	499	61.2	3	71319 - 71378	MOY_35728		

[a]centimorgan
[b]linkage group
[c]single nucleotide polymorphism

Table 4.16: List of 'crumbly' microarray probes matching the two selecting markers (genes target) highly associated with the most significant markers for the two 'crumbly' QTLs on linkage group 3, the one identified previously by Graham *et al*., (2015) and that recently identified in this work.

List of the 'crumbly' microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') and mapped close to the most significant genotype marker of the two 'crumbly' QTLs on linkage group 3. For each probe was reported the scaffold to which the probe matched, the probe position both in cM and in base pair along the scaffold. Moreover, the position of the Glen Moy marker (gene target) along the scaffold and all its SNPs with their position in base pair were reported too.

Pedigree				
♀ mother	**x**	**♂ father**	**genotype**	**phenotype**
complex hybrid	**x**	complex hybrid	Autumn Bliss	'crumbly'
Joan Squire	**x**	complex	Brice	'crumbly'
[1]Autumn bliss			Erika	'crumbly'
SCRI 6531/84	**x**	SCRI 6549/1	Glen Prosen	'crumbly'
			Kweli	'crumbly'
Nootka	**x**	Glen Prosen	Tulameen	'crumbly'
			Obbard	'crumbly'
Willamette	**x**	Cuthbert	Meeker	'crumbly'
97134B1	**x**	8510A57	0867E-4	'crumbly'
7326E1	**x**	7412H16	Glen Rosa	'crumbly'
			Imara	'crumbly'
Glen Rosa	**x**	SCRI 8605C-2	Glen Doll	no 'crumbly'
Preussen	**x**	Lloyd George	Malling Minerva	no 'crumbly'
			Malling Leo	no 'crumbly'

[1]open pollinated

Table 4.17: Selected genotypes for the validation pool that did not give any amplification fragments for all the five SNPs markers, polymorphic between Glen Moy and Latham.

List of fourteen genotypes, out of the 63 selected for the GWAS and not containing the five markers strongly associated with the three 'crumbly' QTLs and then chosen for this study. For each genotype, when available, was reported the pedigree too.

Chi-square (χ^2) tests of independence were used to look for associations between the two 'crumbly' categories (i.e. 'crumbly' and non 'crumbly') and markers. Each marker SNP position was tested to see whether any of the various polymorphic forms detected in the GWAS population was significantly associated with either of the two traits (i.e. 'crumbly'

and non ‘crumbly’), at 95% confidence levels. If this was found then that SNP position could have been considered linked with the genotypes showing ‘crumbly’ phenotype and that marker could be then analysed by a generalised mixed model (Yu et al., 2006) specifically designed to address quantitative traits, such as ‘crumbly’ fruit and complex levels of relatedness within the analysed population.

None of the segregating SNP position, for all the five markers, was significatively associated with the ‘crumbly’ phenotype in the studied population (see Tables 4.18 and 4.19 for further details). The use only of the forward primers to sequence the amplified fragment containing the selected limited the size of the region which was screened for segregating alleles. The use of both forward and reverse primers and the chromosome walking, along the region where the markers are located, could increase the size of the fragment to be sequenced and thus the number of SNP positions to enhance the chances of identification of segregating SNPs, potentially associated with the ‘crumbly’ phenotype, in the selected population.

The presence of expected values lower than 5 in any cell of the contingency table would make the probability test calculation unreliable (Steve, 2011). To bypass this issue, a permutation test to calculate the significance probability for a chi-square test of the independence of rows and columns was performed. This procedure converts the usual chi-square test in a nonparametric alternative that, in situations where some values in the contingency table are lower than five, allows to strengthen the results of the analysis (GenStat manual, VSN International, UK). The permutation test simulates the random distribution of table values that may occur in tables that have the same overall distribution of numbers over the columns, and over the rows, as in the original table. Even by means of a more accurate and precise statistical analysis, none of the markers tested proved to be significantly associated with the ‘crumbly’ phenotype.

marker	allele	*χ^2	[l]d.f.	p-value
MOY_34151	base_146	1.99	3	0.574
	base_139	2.62	4	0.662
	base_183	0.88	2	0.643
	base_158	2.56	2	0.217
[a]MOY_35	base_75	2.91	4	0.573
	base_174	2.52	2	0.284
	base_242	0.68	1	0.41
s182_p9185_R6	base_110	1.96	3	0.58
	base_129	2.93	2	0.231
	base_139	6.39	4	0.172
	base_147	1.52	2	0.467
	base_148	2.18	3	0.536
	base_170	3.19	2	0.203
	base_206	4.11	3	0.250
	base_236	2.01	2	0.367
	base_376	1.14	2	0.565
	base_441	2.18	3	0.536
	base_444	2.71	2	0.257
[b]s353_p21	base_196	3.69	3	0.297
	base_180	0.94	2	0.625
	base_303	2.9	2	0.235

*Pearson chi-square value
[l]degree of freedom
[a]**MOY_35728**
[b]**s353_p21288_R16**

Table 4.18: Significance probability for a chi-square test .

For each marker was reported the number of alleles, the chi-square test statistic, the probability test (p-value) and the number of degrees of freedom for a chi-square test with two criteria (i.e. 'crumbly' and 'non crumbly') was calculated according the formula (m-1) x (n-1) where m is the number or rows (i.e. the two different phenotypes, 'crumbly' and non 'crumbly') and n the number of column (i.e. the number of different alleles) of the contingency table.

marker	allele	*χ^2	[l]d.f.	p-value
MOY_34151	base_146	1.99	3	0.668
	base_139	2.62	4	0.736
	base_183	0.88	2	0.723
	base_158	2.56	2	0.343
[a]MOY_35	base_75	2.91	4	0.684
	base_174	2.52	2	0.379
	base_242	0.68	1	0.479
s182_p9185_R6	base_110	1.96	3	0.727
	base_129	2.93	2	0.256
	base_139	6.39	4	0.167
	base_147	1.52	2	0.612
	base_148	2.18	3	0.632
	base_170	3.19	2	0.240
	base_206	4.11	3	0.234
	base_236	2.01	2	0.367
	base_376	1.14	2	0.569
	base_441	2.18	3	0.609
	base_444	2.71	2	0.298
[b]s353_p21	base_196	3.69	3	0.346
	base_180	0.94	2	1
	base_303	2.9	2	0.242

*Pearson chi-square value
[l]degree of freedom
[a]**MOY_35728**
[b]**s353_p21288_R16**

Table 4.19: Significance probability for a for a permutated chi-square test.

For each marker was reported the number of alleles, the chi-square test statistic, the probability test (p-value) and the number of degrees of freedom for a chi-square test with two criteria (i.e. 'crumbly' and 'non crumbly') was calculated according the formula (m-1) x (n-1) where m is the number or rows (i.e. the two different phenotypes, 'crumbly' and non 'crumbly') and n the number of column (i.e. the number of different alleles) of the contingency table.

4.4 Discussion

One hundred and sixty-five probes differently expressed between the two phenotypes, mostly and never 'crumbly' were mapped within the three crumbly QTLs. Two of the QTLs were previously identified (Graham et al., 2015) on linkage groups 1 and 3 and a new QTL was found on linkage group 3 during this work. Of all these probes, seventeen were selected since they matched predicted *R. idaeus* gene whose *A. thaliana* orthologs genes have ontology annotation related to hormones, pollen or flower development which, as for chapter three, were considered within the main potential factors responsible for the setting and growing of 'crumbly' like misshapen fruits. Though other genes and mechanisms could not be discounted it was necessary to focus down to start understanding the control of the condition.

Of particular interest, from a 'crumbly' fruit perspective, was the probe CUST_24407_PI426541283 that was mapped in the original 'crumbly' QTL found on LG3. This probe matched the predicted *R. idaeus* gene whose *A. thaliana* ortholog, AT5G12210.1, encodes RGTB1 (Rab geranylgeranyl transferase subunit 1). Partial deficiency in the expression level of this gene, in *Arabidopsis thaliana*, negatively affected polar growth of pollen tubes compromising the fertilization of the ovules (Gutkowska et al., 2015). Pollen tubes are polar tubular outgrowths of the pollen grains. They grow only at their tips where Rabs proteins are abundant and so it is not surprising that defects in the assembly, function and synthesis of these enzymes reduce cell polarity compromising pollen tube elongation (Gutkowska et al., 2015). Rab geranylgeranyl transferase (RGT) is a heterodimeric enzyme formed by two subunits, α (RGTA) and β (RGTB); in Arabidopsis the β-subunit is encoded by two different genes, AtRGTB1 at locus AT5G12210 and AtRGTB2, at locus AT3G12070 (Hala et al., 2010). Single mutants of the genes coding the β-subunit of the RGT enzyme, *atrgtb1* and *atrgtb2* both produce stems with multiple flowers suggesting a loss of apical dominance. More interesting, even though the number of flower organs is constant, they never fully open and they have protuberant pistils, much longer than the stamens which prevents pollination from the same flower, but only from those adjacent. This adjacent pollination is quite rare (Gutkowska et al., 2015, Hala et al., 2010).

In the crumbly microarray experiment, the probe CUST_24407_PI426541283, differently expressed between the two phenotypes, mostly and never 'crumbly', at only open flower stage, with about 99.9% confidence level (see Table 4.8 for details), matched

the predicted *Arabidopsis thaliana* ortholog gene (AT5G12210.1) which encodes for the β-subunit (RGTB1) of the RGT enzyme. The probe was downregulated in the mostly 'crumbly' phenotype suggesting that those plants may act as the *A. thaliana* single mutant *atrgb1* described by Hala et al., (2010) and Gutkowska et al., (2015). Then the formation of misshapen 'crumbly' like fruits might be explained by the combination of anatomical and molecular defects. The mostly 'crumbly' plants, under-expressing the raspberry equivalent AT5G12210.1 gene might develop flowers with similar characteristic to the *A. thaliana* atrgtb1 mutants, resulting in protruded carpels that cannot be pollinated by the pollen from the anther of the same flower, reducing the chances for the carpels to be fertilised for spatial issues. This phenomenon would be more marked mainly in cases of lack of pollinators, in fact in normal flowers, self-pollination should prevail over that from distant flowers regardless of whether they are from the same or another plant (Gutkowska et al. 2015; Hala et al. 2010). Furthermore, in case the pollen would reach the carpels, in the mostly 'crumbly' plants, the under expression of the raspberry equivalent gene to AT5G12210.1 would impair the ovule fertilization, as in the *A. thaliana* atrgb1 mutants of Gutkowska et al. (2015), since the pollen tube could not elongate sufficiently to reach the ovary.

In summary it might be reasonable to assume that the combination of anatomical (carpels protruded from the stamens) and molecular (pollen tube growth blocked due to reduced cell polarity as a consequence of lack of RAB proteins in the cells tip) disruptions might be responsible for the formation of fruit with a much lower number of drupelets that do not adhere perfectly to each other and so when picked they crumble.

On the QTL, identified during this work on LG3, CUST_16314_PI426541283 was another interesting probe. It was differently expressed between the two phenotypes, mostly and never 'crumbly' (see Table 4.9 for further details).The probes was upregulated in the mostly 'crumbly' plants and differences were statistically significative at both closed bud and green berry stages but not at open flower. The probe matched a predicted *Arabidopsis thaliana* ortholog gene (AT3G55400.1) that codes for a *methionyl-tRNA synthetase* called OVA1 (ovule abortion 1); disruption of this enzyme in mitochondria causes ovule abortion. *Aminoacyl-tRNA synthetases* (AARSs) are essential enzymes catalysing the reaction responsible for the attachment of amino acids to their corresponding tRNAs. In *Arabidopsis thaliana*, Berg et al., (2005) with insertional mutagenesis, produced T-DNA insertion mutants in nine AARSs gene, OVA1 included. These mutants completely lacked the gene function with the resulting plants showing

siliques containing aborted ovules and/or mutant seed arrested soon after fertilization; showing evidence that these AARs gene are required for the development of the embryo. The probe CUST_16314_PI426541283 matched the predicted *A. thaliana* ortholog gene AT3G55400.1 and, although the expression level was low, it was upregulated in the mostly 'crumbly' compared with the never 'crumbly' phenotype, in all the three stages. The differences in the expression levels were significant at about 99.9% of confidence levels for the interaction phenotype/stage at green berry only (see Table 4.9).

The samples at the green stage were collected at the emergence of the drupelets when it might be expected that the embryos were still developing inside the newly forming seeds; it might be reasonable to hypothesize that in the mostly 'crumbly' phenotype the reduced number of drupelets in the fruits could be explained as a consequence of ovule abortion or arrest of seed formation soon after fertilization caused by the upregulation of the *R. idaeus* gene that corresponds to the *A. thaliana* AT3G55400.1 (OVA1), which codes for the aforementioned *methionyl-tRNA synthetase*. A scenario opposite to the one found by Berg at al. (2015) in their experiment on *Arabidopsis thaliana* T-DNA insertion mutants bearing siliques with aborted ovules as previously described.

The results of the crumbly microarray experiment showed an increase of the expression level, for the *R. idaeus* gene corresponding to the *A. thaliana* (OVA1), in the mostly 'crumbly' phenotype; the plant type bearing misshapen fruit with reduced number of drupelets. According to the 'crumbly' microarray results, a simple model based on the downregulation of the *Rubus idaeus*, equivalent gene to the *Arabidopsis thaliana* OVA1, could not explain the formation of misshapen fruit with reduced number of drupelets and in fact, in the never 'crumbly' plants (i.e. plants bearing always normal fruits) the probe CUST_16314_PI426541283 was downregulated but the plants always bore normal shape fruits. An explanation to this apparently contradictory effect of the OVA1 *R. idaeus* equivalent gene could be the complexity behind processes such as ovule fertilization and seed formation that ought to involve many other regulators, probably not even detected with the microarray experiment, acting both up and downstream to OVA1, whose action contributes to radically affect its function.

The probe CUST_54460_PI426541283, mapped inside the original crumbly QTL on LG3 (Graham et al., 2015), and matched the predicted *Arabidopsis thaliana* gene AT1G62360.1 encoding for the KNOX transcription factor SHOOT MERISTEMLESS (STM) which is responsible for promoting carpel development and their associated placental tissues (Scofield et al., 2007). In *Arabidopsis thaliana,* experiments with RNA

interference on the STM gene (STM-RNAi) showed the formation of anatomical defects at the level of the floral meristem. In case of weak STM-RNAi, flowers did not develop proper carpels and they produced less ovules while in case of severe STM-RNA interference, flowers simply developed sepals and stamens lacking carpels completely (Scofield et al., 2007). In the crumbly microarray, the probe CUST_54460_PI426541283 was differentially expressed between the two phenotypes mostly and never 'crumbly' with the difference being significant at about 99.9% of confidence level in closed bud and green berry but not in open flower (see Table 4.8). The closed bud stage, from the perspective of the flower organ development was the main stage of interest in regard to the expression of the *Rubus idaeus* equivalent gene to the *Arabidopsis thaliana* (AT1G62360.1) and in the mostly 'crumbly' phenotype, the expression level was higher than in the never 'crumbly' counterpart. Situations where plants would bear aberrant flowers with reduced number of carpels could easily explain the formation of 'crumbly' like misshapen fruit with lower number of drupelets; this is in line with the *A. thaliana* mutants of the STM-RNAi experiments performed by Scofield et al. (2007). It then could be proposed that in *R. idaeus* too, plants showing a reduced expression of the STM gene, could have flowers with reduced number of carpels and ovules, potentially giving rise to fruits with lower number of drupelets. The microarray however showed an increase of the expression level of the probe matching the predicted *Arabidopsis thaliana* STM gene, in the mostly 'crumbly' phenotype; the one having misshapen 'crumbly' like fruits. This result is contrary to the hypothesis explained by a downregulation of the STM gene; in a similar fashion of what was registered in the in the STM-RNAi mutants by Scofield et al. (2007) and therefore shows another or more complex mechanism for the formation of malformed fruits.

The model for the role of STM in carpel development proposed by Scofield et al. (2007) is more complex and involves another key regulator of flower organ development, the AGAMUS like protein (AG) and its repressors, with BELLRINGER being one of them. In the 'crumbly' microarray the two probes, CUST_20859_PI426541283 and CUST_49079_PI426541283, matching the predicted ortholog *Arabidopsis thaliana* genes coding respectively for AG (AT3G58780.3) and BELLRINGER (AT5G02030.1) were not significantly different with respect to interaction between stages (i.e. closed bud, open flower and green berry) and type (i.e. mostly and never 'crumbly').

Both AG and STM are required for allowing the differentiation, in the floral meristem (FM), of stem cells into carpels (Scofield et al. 2007). The results of the crumbly

microarray showed that STM was upregulated in the mostly 'crumbly' phenotype at both closed bud and green berry stages, with the differences in the expression levels between the two phenotypes being significant at about 99.9% of confidence levels (see Table 4.8). The other player (AGAMUS), in the carpel development model proposed by Scofield et al. (2007), was not differentially expressed between the two phenotypes and the same was recorded for BELLRINGER a transcription factor that repress the expression of AGAMUS. Then, in the mostly 'crumbly' phenotype, the formation of misshapen fruits cannot be explained with the reduction of the expression level of STM and AG predicted genes due to high expression levels of the BELLRINGER (AG repressor). It more reasonable to think that an appropriate balance of all these three factors is necessary for the correct differentiation of the flower regardless the expression levels of STM. A further explanation could be that in *R. idaeus*, the carpels development model would involve other regulatory factors rather than AGAMUS whose probes were not comprised within the 'crumbly' QTLs or the ≈ 56,000 designed for the 'crumbly' microarray.

The re-analysis of the 'crumbly' phenotypic scores with the new GbS (Genotype by Sequencing) linkage map (Hackett et al., 2018) allowed, the identification of a new 'crumbly' QTL on linkage group 3 in this work having its most significant marker at about 62 cM. Together with the more accurate mapping of the two previously identified QTL (Graham et al., 2015). The re-analysis of the fruit ripening scores with the new GbS linkage map also allowed a more accurate location for these QTLs.

From the 'crumbly' fruit perspective, the co-location of crumbly QTLs with those involved in fruit ripening supported earlier findings where rate of ripening has an effect on prevalence of 'crumbly' fruit (see Tables 4.9 and 4.10 for further details). In particular, the association of the original 'crumbly' QTL, identified by Graham et al., (2015), with three fruit ripening QTLs (i.e. PCO2, fruit set and green fruit) representing the early stages in fruit development highlighted the potential influence on both these phenotypes (i.e. 'crumbly' and fruit ripening) of five genes that were differentially expressed with 99.9% confidence level in the 'crumbly' microarray experiment . Of these five *Arabidopsis thaliana* ortholog genes (see Table 4.12) two, AT5G12210.1 and AT1G62360.1, were previously described in this paragraph for their potential role in the 'crumbly' fruit. In their work, Graham et al., (2015) found that the association between crumbly and fruit ripening at fruit set and green fruit level was positive with the longer the fruit took to reach fruit set and green fruit stage the more likely to be 'crumbly'. The results of the 'crumbly' microarray experiment, showed that the downregulation of the ortholog

AT5G12210.1 might be linked with the formation of misshapen flowers with protuberant carpels whose pollination by from the stamen of the same flowers could be negatively affected.

The re-analysis of the 'crumbly' phenotypic scores (Graham et al., 2015) by means of more powerful technologies such as the new GbS linkage map (Hackett et al., 2018), allowed the identification of a new 'crumbly' QTL and of new and more significative markers for the other two crumbly QTLs previously identified by Graham et al., (2015). The logical continuation for such work would be the identification and selection of one or more markers, suitable for commercial applications. Such marker/s are highly welcomed by the industry because they represent a very powerful tool to help tackling problems such as 'crumbly' fruit. They can be used in molecular assisted breeding for selection of potential new varieties 'crumbly' free but can be even used in diagnostics to assess the health of the mother stock plants and help reducing the propagation of material carrying the 'crumbly' markers and then potentially more susceptible.

The first 'crumbly' markers tested for association with specific genotypes from a population of 63 individuals did not show any significant associations with the 'crumbly' fruit phenotype. Due to time constraints the first steps towards marker identification were focused on the sequencing of amplified regions produced with only the forward primers. This produced a limited number of SNPs for analysis and in future the use of both primers (i.e. forward and reverse) and extending sequencing across a wider region might help with the identification of further alleles that might be significantly associated with 'crumbly' fruit. Three more 'crumbly' markers, SSR type, have been selected and their primers have been designed. The amplified regions have been prepared for the fragment analysis whose results should allow to verify whether at least one of these three markers could be significatively associated with 'crumbly' fruit. Unfortunately, so far, the lock down imposed by the March 2020 corona virus pandemic prevented the analysis being completed. The selection of markers associated to complex quantitative traits such as 'crumbly' fruit is not an easy task and the size and structure of the population taken under exam play an important role in the whole process.

4.5 Conclusions

The Glen Moy x Latham raspberry linkage map, developed using Genotype by Sequencing (Hackett et al., 2018), allowed greater understanding of the 'crumbly' fruit phenomenon. First of all, a new 'crumbly' QTL on linkage group 3 was identified together with the more accurate positioning of the two 'crumbly' QTLs previously found on linkage 1 and 3 (Graham et al., 2015). The GbS map was linked to the genome sequences of Glen Moy and a partial Latham sequence. The study of the function of the genes mapped inside these QTLs showed interesting molecular processes behind 'crumbly' fruit. Moreover, the identification of robust molecular markers strongly associated with 'crumbly' fruit paves the road for future study to help develop practical measures to control and potentially eradicate the condition (i.e. markers assisted breeding).

The study of the expression levels of the probes mapped inside the three 'crumbly' QTL identified similar mechanisms to those highlighted in chapter three. From a molecular/physiological perspective the results of the analysis showed that potential crucial phenomena responsible for the formation of 'crumbly' like misshapen fruits were again anatomical defects of flowers. Examples can be protruded carpels that cannot be self-pollinated due to spatial constraints or flowers with reduced number of carpels that then are suspected to give rise to fruit with reduced number of drupelets that do not adhere perfectly to each other causing the fruit to crumble when it is picked up. Again, as in chapter three, pollen tube that cannot elongate enough to reach the ovule and then fertilize the egg cell or even ovule abortion are all processes that might be taken in great consideration as triggers of crumbly fruit. All these results were in line with those found in chapter three strengthening these discoveries because they represent validation of the results found in chapter three when the same analysis was conducted but in this case on the microarray probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') that specifically mapped inside the three 'crumbly' QTLs.

The new GbS linkage map allowed an additional QTL for crumbly fruit to be identified and also allowed better definition of the relationship between the 'crumbly' phenotype and fruit ripening QTLs. Of interest was the 'crumbly' QTL on linkage group 3, the one identified previously (Graham et al. (2015)). This overlapped three fruit ripening QTLs (see section 4.3.5) but of particular interest was the QTL related to fruit set. In their work Graham et al. (2015) found positive correlations between 'crumbly' fruit and time to

reach fruit set. As the time to reach fruit set protracts, the higher probability that the fruit could be misshapen ‘crumbly’ like.

Last but not least, the new and more significant markers, for all the three ‘crumbly’ QTLs, here identified were used to assess association with ‘crumbly’ fruit in a validation pool. In the first set of SNPs identified, no markers were found to be associated with either phenotype within the population of 63 distantly related genotypes.

Chapter 5: Hormones profiling in artificially induced ‘crumbly’ fruit at two different development stages, green and red berry.

5.1 Introduction

Plants have evolved many different strategies to cope with the challenging environmental conditions they encounter (Wasternack and Kombrink, 2010). Plants cannot move to escape stress factors whether abiotic or biotic, and for this reason they have evolved a series of traits that allow them, for instance, to regenerate damaged organs and tissues and/or to re-direct growth in response to external stimuli (Vanneste and Friml, 2009). Small endogenous signalling molecules, called phytohormones (plant hormones), are responsible for the control and coordination of the physiological processes that plants activate to react to external environmental factors, as well as to regulate their growth (Santner et al., 2009, Wasternack and Kombrink, 2010). Such compounds act locally at very low concentrations, at or near the site of synthesis or even in distant tissues (Santner et al., 2009). In general plant hormones act through extensive crosstalk between themselves and/or other signalling pathways and the results of these interactions are additive, synergistic or antagonistic actions that determine specific and complex physiological outcomes (Vanneste and Friml, 2009, Pan et al., 2008).

Phytohormones, according to their structure and function, are divided into nine classes: abscisic acid (**ABA**), auxins (**AUXs**), cytokinins (**CKs**), gibberellins (**GAs**), jasmonates (**JAs**), salicylates (**SAs**), brassinosteroids (**BRs**), strigolactones (**SLs**) and ethylene (Santner et al., 2009, Zwanenburg et al., 2016). Although each class of phytohormones is linked with specific and typical biological effects (e.g. SA for plant defence, GAs for organ elongation and flowering time, CK_s for germination, etc.), the biological processes are regulated by complex networks involving different hormonal signals (Cao et al., 2016b). In plants, hormone interactions can occur at least at two different levels, hormones distribution (i.e. the opposing action of **AUX_s** and **CK_s** during lateral root initiation) and gene expression (i.e. **AUXs** and **BRs** repress the same genes suggesting coordination between the two signalling pathways) (Santner et al., 2009). The activity of plant hormones depends on their availability which is in turn affected by their metabolism, localisation, transport and signal transduction; the modulation at any of these levels, and there are myriads of possible combinations, can determine different physiological processes (Simura et al., 2018).

To date, analytical methods designed to simultaneously analyse as many classes of phytohormones as possible, are limited. In the last two/three decades, phytohormones have been extensively studied and many protocols have been examined to develop the

analytical methods to detect hormones content in plants. In particular great technological progress has been achieved with immunoassays (Weiler, 1984) and hyphenated techniques such as Gas Chromatography-Mass Spectrometry (GC-MS) (Muller et al., 2002, Kowalczyk and Sandberg, 2001) and Liquid Chromatography-Mass Spectrometry (LC-MS) (Simura et al., 2018, Trapp et al., 2014, Pan et al., 2008, Cao et al., 2016b). The immunological detection methods are characterized by high sensitivity and because they do not require multi step extraction procedures, no loss of material is experienced with the analysis but obviously they are very specific, requiring a method for each hormone to be detected. For this reason they are unsuitable and too laborious for the simultaneous detection of multiple classes of phytohormones (Trapp et al., 2014). The GC-MS methods have been extensively used in the past and are still the most employed analytical system for the simultaneous detection of multiple hormones but the derivatization steps that must be applied to enhance compounds volatility, to improve their stability and facilitate GC separation and MS ionisation (Du et al., 2012b) make the analysis laborious and the high temperatures reached in the GC injector could easily degrade and transform the target analytes, compromising the method reliability (Cao et al., 2016b).

LC-MS, especially in the last decade, has become a valid alternative technique to trace plant hormones as it can overcome the drawbacks of the GC-MS and higher sensitivity/lower detection limits that can be reached with these systems (Cao et al., 2016b). There are many different LC-MS techniques available for accurate analyte quantification but LC-triple quadrupole mass spectrometry, although featuring lower resolution compared to the time-of-flight-MS or the ability to screen for unknown compounds like the orbitrap-MS, represents the best choice when it comes to detection/quantification of target compounds by virtue of its greater sensitivity, repeatability and a wider dynamic range (Cao et al., 2016b).

The correct analytical instrument is only a small part of a complete separation based analytical process. Sample preparation is crucial and can account for the 80% of the total time of analysis. All the following steps of identification, confirmation and quantification can be greatly affected by the extraction process, being crucial in the analysis of trace compounds such as phytohormones (Du et al., 2012b). In general sample preparation involves many procedures: sampling, freeze drying, comminution (i.e. fine grinding with mortar and pestle), homogenization, extraction from the matrix and purification. This last procedure being very important, especially for complex plant extraction, where,

whichever instrument is selected, is not recommended to load the crude plant extract directly through the column to avoid its damage and fouling (Du et al., 2012b) .

Many methods and commercial products have been developed for the removal of the sample matrix and the enrichment of the target molecules (Cao et al., 2016b). Liquid-liquid extraction (LLE) exploits the difference in solubility of the target compounds between two liquid phases; for instance, by separating analytes from an aqueous solution to a non-polar or less polar solvent (i.e. dichloromethane/isopropanol, hexane, diethyl ether, etc.).

Solid phase extraction (SPE), on the other hand, exploits the interaction (i.e. absorption, hydrogen bonding, polar and non-polar interactions, cation and anion exchange and size exclusion) between the analytes and the sorbent material, knowledge of the optimum mechanism of interaction helps in the choice of the most suitable and best performing sorbent material thereby increasing the purification efficiency (Du et al., 2012b).

One of the main challenges for future work is the network plus crosstalk of the hormonal circuits underlying the whole fruit development process. An analytical method designed to detect six groups of plant hormones (i.e. auxins, cytokinins, gibberellins, jasmonates, salicylates and abscisic acid), although only partially covering the whole spectrum of plant hormones (nine classes in total), represents a valid starting point to help understand the metabolic pathways that regulate fruit growth in raspberry undergoing normal developmental processes compared with fruits induced to express the ‘crumbly’ fruit phenotype by mechanical damage to the flowers before pollination. The majority of data available in the scientific literature indicate the involvement of different hormones in the regulation of fruit development. Although the concerted action of auxins and/or gibberellins and/or cytokinins, through their biosynthesis and/or signalling pathway, seems to play a major role, on the basis of the high complexity of the whole process, the involvement of other phytohormones cannot be excluded a priori. In the raspberry ‘crumbly’ fruit, where berry abnormalities occur, different hormonal pathways might be involved in the regulation of these misshapen fruits and then a larger spectrum of different plant hormones must be taken into account to better understand the whole process. An analytical method designed to detect eighteen different plant hormones, covering six groups represents a good starting point to help understand the hormonal molecular circuits behind the fruit development.

In this work, a simple analytical method was developed to detect 18 phytohormones representing six of the most important classes of plant hormones: abscisic acid (**ABA**), salicylates (benzoic acid **BA**, cinnamic acid **CA**, salicylic acid **SA**, methyl benzoate **MeBA**, methyl cinnamate **MeCA** and methyl salicylate **MeSA**), jasmonates (jasmonic acid **JA**, methyl jasmonate **MeJA** and 13-epi-12-oxo-phytodienoic acid **OPDA**), auxins (indole-3-acetic acid **IAA**, indole-3-butyric acid **IBA**, Indole-3-carboxylic acid **ICA** and methyl indole-3-acetate **MeIAA**), gibberellins (gibberellic acid A_1 **GA_1**, gibberellic acid A_3 **GA_3**, gibberellic acid A_4 **GA_4**) and cytokinin (**zeatin**). The method was developed to allow the analysis of the raspberry fruit extract at two different stages, green berry and red berry both in the receptacle and in the drupelets to try to verify the hypothesis that fruit development is coordinated and regulated by the receptacle acting as a leading hub, synchronizing the growth of the many fertilized ovaries that will constitute the fruit drupelets. Samples of 'crumbly' induced fruit (flowers mechanically damaged) and control (normal developed fruits) were collected. Drupelets and receptacle for each berry were separated with the aim of verifying which phytohormones were primarily involved in the fruit development and what were, if any, the differences between artificial 'crumbly' fruit and a normal fruit (control) and how, if the hypothesis was correct, the receptacle fulfilled its function of raspberry fruit regulator.

5.2 Materials and methods

5.2.1 'Crumbly' fruit induction experiments

Glen Ample long canes were purchased from EU plants Ltd. (Abingdon, UK). Plants were cultured in controlled environment, plant growth room, (Nijssen, The Netherlands) and for the first two weeks the following parameters were set: minimum temperature 10 °C, maximum temperature 14 °C, relative humidity 70% and daylight length 16 hours (from 7:00 AM to 11:00 PM); such conditions helped the plants to acclimatize. After the two weeks of acclimatization, all the environmental parameters remained unchanged except the maximum temperature that was increased by four degrees, from 14 to 18 °C. The first flowers, immediately after anthesis, were collected and dried at room temperature for two days to assist in pollen maturity. The dried flowers were closed in Petri dishes and stored in the fridge; they were used as pollen reservoir for the 'crumbly' induction experiments. Once the plants were established under these conditions, all unopened flowers one day from anthesis were emasculated. The emasculations were performed according to the method available at The James Hutton Institute (N. Jennings, James Hutton Ltd. Raspberry breeder pers. comm.). All flower buds were emasculated by means of a scalpel blade by simply following the contour of the five sepals, in this manner: petals, sepal and stamens were cut away. After the two days required for the stigmas to become receptive to pollen, the emasculated flowers were first damaged and then hand pollinated. The damage, carried out on the receptacle, was accomplished by pinning its tip and side with a needle and for each damaged flower (treatment) a control was carried out by simply hand pollinating the emasculated flower. Each treated bud (damaged receptacle and control) was tagged reporting the date of the emasculation, the kind of treatment (damaged receptacle or control) carried out and the number of the plant.

Fruits derived from the 'crumbly' fruit induction experiment (previously described in this work, i.e. damaged receptacle and control), were collected at two different stages, green and re berry and immediately frozen in liquid nitrogen. Obviously, part of the samples were artificially induced 'crumbly', in this case of those derived from the damage of the receptacle, while the remaining samples (i.e. the control) were normal shape fruit. All these samples were then stored at -80 °C to be used for the hormone profiling analysis.

5.2.2 Chemicals and materials

The 18 authentic standards of phytohormones were purchased from three different suppliers. SigmaAldrich Co. Ltd. (Gillingham, UK): IAA, IBA, ICA, MeIAA, SA, CA, BA, MeBA, MeCA, MeSA, GA_3, GA_4, and MeJA. OPDA and JA were purchased from Cayman Chemicals (Cambridge bioscience, Cambridge, UK) while GA_1 was purchased from OlchemIm Ltd (Olomouc, Czech Republic). All the organic solvents used for the extraction (methanol, acetonitrile, 2-propanol, HCl and dichloromethane) were analytical grade and were purchased from Fisher-Scientific (Loughborough, UK). The chromatographic separation was performed with formic acid, methanol and water HPLC grade (these last two supplied by J. T. Baker) that were purchased from Fisher-Scientific Ltd (Loughborough, UK). Deionised water, used for all the aqueous solutions, was obtained using ELGA DV25 system (High Wycombe, UK). Isotopically labelled standards (trans cinnamic-d_7 acid, indole-3-acetic-2,2-d_2 and Jasmonic-d_5-acid) were purchased from SigmaAldrich Co. Ltd. (Gillingham, UK) while ([$^2H_2$]GA_1, [$^2H_2$]GA_3, [$^2H_5$]MeIAA, [$^2H_6$]-trans zeatin, [$^2H_6$](+)-cis,trans-ABA and [$^2H_4$]SA) were purchased from OlchemIm Ltd. (Olomouc, Czech Republic).

5.2.3 Preparation of standard mix solutions

For each target phytohormone, a stock solution was prepared at part per million (ppm) concentration in pure methanol or pure acetonitrile or pure distilled water, depending on the compound solubility (Table 5.1). A standard mix solution was prepared diluting the stock solution in CH_3OH:H_2O (50:50) and the final concentration of each compound was empirically optimized by analysing progressively decreasing standards mix concentrations; the compounds that ionize better are better detected by the instrument and are obviously required in lower concentration. In Table 5.1 the final concentration of each target plant hormone in the final standards mix solution was identified. For the nine isotopically labelled standards, the mix solution was prepared by diluting the stock solutions at different concentrations depending on the range of the calibration curve; the final concentration for each of the deuterated standard was reported in Table 5.2.

Compound	solvent	Stock solution [ppm]	Standard mix [ppm]
ABA	methanol	3250	12
SA	methanol	2800	20
JA	ethanol	10000	2
CA	methanol	4950	4
GA_4	methanol	5000	30
IAA	methanol	5000	1
IBA	methanol	5800	3
ICA	methanol	4400	30
MeBA	water	2092	8
MeCA	methanol	3350	20
MeIAA	methanol	3250	4
MeJA	methanol	5150	1
MeSA	water	587	30
OPDA	ethanol	1000000	2000
Zeatin	ethanol	500	3
GA_3	acetonitrile	4600	2
GA_1	methanol	1000	4
BA	methanol	5000	2

Table 5.1: Target compounds stock and standards mix solutions and their concentration in the standards mix solution (2 mL final volume).

For each compound the concentration, in part per million (ppm) of the stock solution (2 mL total volume) and the solvent in which the compounds were suspended was reported. The final concentration of each compound in the standards mix solution containing all the 18 compounds was reported in the last column of the table.

Compound	solvent	Stock solution [ppm]	Standard mix [ppm]
d-SA [$^2H_4$]	methanol	5	0.15
d-MeIAA [$^2H_5$]	methanol	5	0.015
d-IAA [$^2H_2$]	methanol	0.5	0.005
d-CA [$^2H_7$]	methanol	5	0.06
d-Zeatin [$^2H_5$]	methanol	1	0.08
d-GA_1 [$^2H_2$]	methanol	0.5	0.03
d-GA_3 [$^2H_2$]	methanol	0.5	0.018
d-ABA [$^2H_6$]	methanol	0.5	0.08
d-JA [$^2H_5$]	methanol	0.5	0.015

Table 5.2: Isotopically labelled internal standards (ISTD) stock and standards mix solutions.
For each ISTD the concentration, in part per million (ppm), of the stock solution (5 mL total volume) and the solvent in which the compounds were suspended was reported. The final concentration of each labelled compound in the ISTD mix solution containing all the 9 compounds was reported in the last column of the table.

5.2.4 Phytohormone extraction

Plant material (both drupelets and receptacle from both green and red stages) was first ground to a fine powder in liquid nitrogen with mortar and pestle. The same procedure was repeated for both the treatment samples (i.e. control and damaged receptacle). The sample was then weighed (50 mg) into 2 mL Eppendorf tube for its homogenization. Inside the tube, 2 stainless steel beads (3 mm of diameter) were added together with 1 mL of extraction solution, containing a mixture of isotopically labelled internal standards. The extraction solution, ice-cold 50% aqueous acetonitrile, was spiked with nine isotopically labelled standards ([$^2H_7$]trans-CA, [$^2H_2$]IAA, [$^2H_5$]JA, [$^2H_2$]GA_1, [$^2H_2$]GA_3, [$^2H_5$]MeIAA, [$^2H_6$]trans-Zeatin, [$^2H_6$](+)-*cis*, *trans*-ABA and [$^2H_4$]SA), in amounts ranging from 0.015 to 0.15 ppm (part per million). This was done to have at least one internal standard for each of the six phytohormones groups; the best compromise to allow an absolute quantification of the endogenous plant hormones in the samples analysed. The Eppendorf tubes so prepared, were loaded onto a vibration mill, model MM 301

(Retsch GmbH & Co. KG, Haan, Germany) with an operating frequency of 30 Hz for 1 minute and homogenized. The mixture was first sonicated for 3 minutes in a cold room (4 °C) and then extracted using a benchtop laboratory rotator at 15 rpm for 30 min in cold room (4 °C). The extraction was followed by centrifugation at 16,600 g at 4 °C for 10 min, the supernatant was saved and transferred to a new 2 mL Eppendorf tube for further purification. Samples were purified using Oasis® HLB reverse-phase (C_{18}), polymer-based, solid phase extraction (RP-SPE) cartridge (1 cc/30 mg) from Waters Ltd (Elstree, UK). The solid phase extraction protocol consists of four steps. First, the cartridge was conditioned/activated with 1 mL of pure methanol followed by 1 mL of distilled water. The conditioning/activation was followed by the column equilibration step with 1 mL of 50% aqueous acetonitrile (v:v) and then by the sample loading. The flow through was collected in a 7 mL amber glass vial (SigmaAldrich Co. Ltd, UK) together with 1 mL of 30% aqueous acetonitrile solution (v:v) used to rinse the cartridge. Samples were dried in speed vacuum at 30 °C, for about 30 minutes to help remove all the organic solvent, soon after they were frozen in liquid nitrogen and then loaded on a freeze drier (Martin Christ Gefriertrocknungsanlagen GmbH, Germany) to be dried overnight. Once dried, the samples were briefly stored at -80 °C; before performing the analysis the samples were reconstituted with 40 µL of 30% aqueous acetonitrile (v:v) and analysed with the LC-MS/MS system.

5.2.5 Chemical analysis

Chemical analysis of plant material (receptacle and drupelets) was performed on an Agilent 1260 high performance liquid chromatography (HPLC) system consisting of quaternary pump, a Diode Array Detector (DAD), a Temperature Control Device and a solvent Thermostat module (Agilent Infinity 1290) coupled to an Agilent 6460A Triple Quadrupole Mass Spectrometer (Agilent Technologies, Santa Clara, CA, USA). The sample extract (5 µL) was injected onto a 150 x 2 mm (5µm) Gemini RP C18 (110 Å) column fitted with a Gemini C18 4 x 2 mm Security guard cartridge (Phenomenex, Cheshire, UK). Samples were eluted at a flow rate of 0.3 ml min^{-1}, using a gradient separation with two mobile phases A = 0.1% formic acid in deionized water and B = 0.1% formic acid in methanol. The elution gradient lasted 29 min in total and was as follows: A/B 95/5 (v/v) hold for 2 min, ramped up to 35% B in 1 min, followed by further ramping up from 35% to 55% in 15.5 min and further ramped up from 55% up to 100% in 3 min

and hold for 2 min. Within 0.5 min the gradient was returned to the initial composition of 5% and held for 5 min (Table 5.3).

Mass detection was carried out in negative or positive ion mode depending on the phytohormone (Tables 5.4 and 5.5) using a jet stream electrospray ionization (ESI) interface coupled to the triple quadrupole system (Agilent, USA). For ESI, the gas temperature, gas flow, nebulizer pressure, sheath gas temperature, sheath gas flow, capillary cap and nozzle voltage were optimized with the help of Agilent Source Optimizer Software and set to 350 °C, 10 L min^{-1}, 30 psi, 4000 °C, 11 L min^{-1} and 3500 V (both ion modes). Collision energies and fragmentor voltages for transition states of the 18 phytohormones were optimized using Agilent Optimizer Software (Agilent Technologies, USA). Hereby the most sensitive transitions (i.e. transitions with the highest intensity of the product ions were chosen to build the final multiple reactions monitoring (MRM) method. As shown in Tables 5.4 and 5.5, 45 transitions were part of the MRM mode, each with a dwell time of 10 ms and a delay time of 3.5 ms, leading to a total cycle time of 445.5 ms and thus 2.5 Hz (2.5 cycles per second). Peak areas of the 18 phytohormones and of the nine isotopically labelled internal standards were integrated with Agilent MassHunter Quantitative Software (Agilent, USA; version B.07.00).

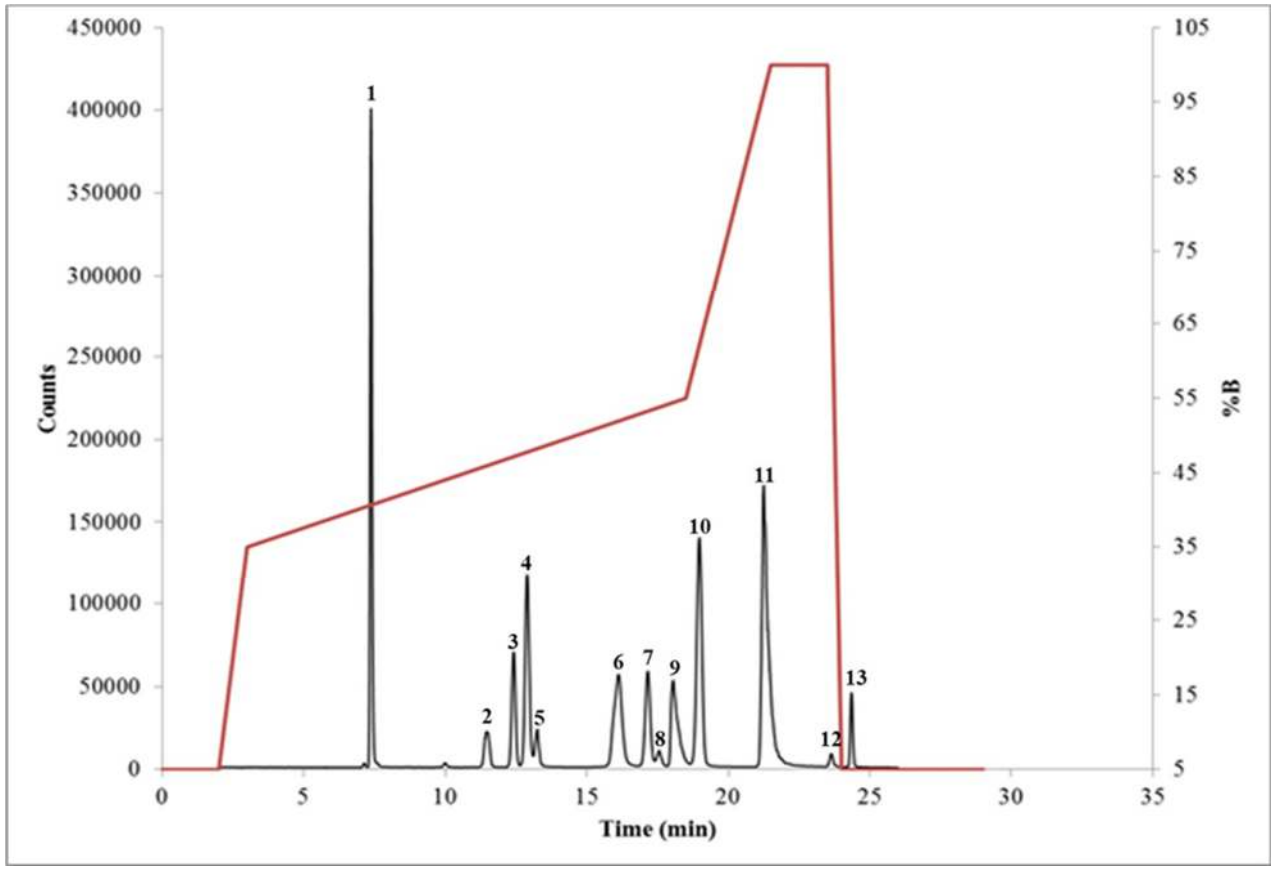


Figure 5.1: Standards mix chromatogram and elution gradient.

Total ions count chromatogram of the standards mix containing the 18 targets compounds and respective applied elution gradient (%B) using Liquid Chromatography-Electrospray Ionisation-Multiple Reaction Monitoring (LC-ESI-MRM). Thirteen out of the total eighteen compounds were clearly displayed in the chromatogram and were labelled each with a number from 1 to 13 that corresponded respectively to: 1 (Zeatin), 2 (GA_3), 3 (GA_1), 4 (ICA), 5 (IAA), 6 (SA), 7 (ABA), 8 (MeIAA), 9 (IBA), 10 (JA), 11 (MeSA), 12 (MeJA) and 13 (GA_4).

Time (min)	A (%)	B (%)	flow (mL min^{-1})
0	95	5	0.3
2	95	5	0.3
3	65	35	0.3
18.5	45	55	0.3
21.5	0	100	0.3
23.5	0	100	0.3
24	95	5	0.3
29	95	5	0.3

Table 5.3: Chromatographic elution gradient.

The chromatographic separation was performed with a flow rate of 0.3 mL per minute. The mobile phase, (elution solution) was a mix of water (A) and methanol (B) in a concentration gradient starting with a 95% of water hold for two minutes, then in 1 minute the methanol concentration was ramped up to 35%, followed by a more gradual ramping up of methanol (B) concentration, from 35% to 55% in 15.5 minutes to finally reach 100% of methanol (B) in 3 minutes. The 100% methanol (B) concentration was hold for two minutes and then quickly (0.5 minutes) returned to 5% and held for 5 minutes.

Compound	[1]prec. ion	[2]prod. ion	dwell	[3]Fragm.	[4]c. e.	polarity	[5]tR
d-GA$_1$ [$^2H_2$]	351.2	333.2	10	105	4	positive	11.68
d-GA$_1$ [$^2H_2$]	351.2	305.2	10	105	4	positive	11.68
GA$_1$	349.2	331.1	10	95	20	positive	11.89
GA$_1$	349.2	285.1	10	95	16	positive	11.89
d-GA$_3$ [$^2H_2$]	347.1	241.2	10	150	12	negative	11.2
d-GA$_3$ [$^2H_2$]	347.1	143.1	10	150	32	negative	11.2
GA$_3$	345.1	239.2	10	155	12	negative	11.44
GA$_3$	345.1	143.1	10	155	24	negative	11.44
GA$_4$	331.1	287.2	10	220	16	negative	24.28
GA$_4$	331.1	257.1	10	220	20	negative	24.28
OPDA	291.2	247.2	10	175	12	negative	25.43
OPDA	291.2	165.1	10	175	16	negative	25.43
d-ABA [$^2H_6$]	269.2	225.2	10	190	8	negative	15.65
d-ABA [$^2H_6$]	269.2	159.1	10	190	8	negative	15.65
ABA	263.1	219.1	10	95	8	negative	16.04
ABA	263.1	153.1	10	95	4	negative	16.04
MeJA	225.1	151.1	10	85	8	positive	23.95
MeJA	225.1	133.1	10	85	12	positive	23.95
***d-Z. [$^2H_5$]**	225.1	137	10	100	16	positive	7.2
***d-Z. [$^2H_5$]**	225.1	135.9	10	100	16	positive	7.2
Zeatin	220.1	202.1	10	90	8	positive	7.32
Zeatin	220.1	136	10	90	16	positive	7.32
d-JA [$^2H_5$]	214.2	62.1	10	130	8	negative	18.95
d-JA [$^2H_5$]	214.2	42.1	10	130	48	negative	18.95
JA	209.1	109	10	113	16	negative	19.36
JA	209.1	59.1	10	113	8	negative	19.36
IBA	204.1	186.1	10	90	8	positive	18.95
IBA	204.1	130.1	10	90	28	positive	18.95
[1]d-Me. [$^2H_5$]	195.1	135	10	90	16	positive	16.6
[1]d-Me. [$^2H_5$]	195.1	134.1	10	90	12	positive	16.6
MeIAA	190.1	130	10	75	12	positive	17.14
MeIAA	190.1	77.1	10	75	50	positive	17.14

[1]precursor ion - [2]product ion - [3]fragmentator - [4]collision energy - [5]retention time (min)
***d-Zeatin [2H5]**
[1]d-MeIAA [$^2H_5$]

Table 5.4: Mass spectrometry optimized parameters for the multiple reactions monitoring (MRM) method (part I).

For each of the 18 target compounds and the 9 isotopically labelled internal standards the best parameters for their detection were reported in the table. These parameters were, the mass of the precursor ion and of its product after fragmentation reaction (product ion), the dwell time (the duration in which each m/z (mass to charge ratio) ion signal is collected) of 10 ms. The fragmentation voltage that controls the speed at which the ion moves from the electrospray chamber to the mass spectrometer and the collision energy that controls the rate of acceleration as the ions enter the quadrupole 2 (Q2) and then regulate fragment ion intensity. The best ionization mode (i.e. positive or negative) and the retention time at which each compound is eluted through the chromatographic column were reported too.

Compound	[1]prec. ion	[2]prod. ion	dwell	[3]Fragm.	[4]c. e.	polarity	[5]tR
d-IAA [$^2H_2$]	178.1	132	10	90	12	positive	12.6
d-IAA [$^2H_2$]	178.1	78.1	10	90	48	positive	12.6
IAA	176.1	130.1	10	85	12	positive	12.89
IAA	176.1	77.1	10	85	50	positive	12.89
MeCA	163.1	131	10	70	8	positive	23.62
MeCA	163.1	103.1	10	70	20	positive	23.62
ICA	162.1	144	10	80	12	positive	12.47
ICA	162.1	118.1	10	80	12	positive	12.47
MeSA	153.1	121	10	85	12	positive	21.48
MeSA	153.1	65.1	10	85	32	positive	21.48
d-CA [$^2H_7$]	154.1	110.1	10	115	12	negative	16.97
MeBA	137.1	100	10	85	4	positive	18.41
MeBA	137.1	77.1	10	85	28	positive	18.41
d-SA [$^2H_4$]	141.1	97	10	90	16	negative	15.1
d-SA [$^2H_4$]	141.1	69.1	10	90	31	negative	15.1
SA	137	93.1	10	75	16	negative	15.49
SA	137	65.1	10	75	31	negative	15.49
BA	121	77.1	10	105	8	negative	12.98

[1]precursor ion - [2]product ion - [3]fragmentator - [4]collision energy - [5]retention time (min)
*__d-Zeatin [2H5]__
[1]__d-MeIAA [$^2H_5$]__

Table 5.5: Mass spectrometry optimized parameters for the multiple reactions monitoring (MRM) method (part II).

For each of the 18 target compounds and the 9 isotopically labelled internal standards the best parameters for their detection were reported in the table. These parameters were, the mass of the precursor ion and of its product after fragmentation reaction (product ion), the dwell time (the duration in which each m/z (mass to charge ratio) ion signal is collected) of 10 ms. The fragmentation voltage that controls the speed at which the ion moves from the electrospray chamber to the mass spectrometer and the collision energy that controls the rate of acceleration as the ions enter the quadrupole 2 (Q2) and then regulate fragment ion intensity. The best ionization mode (i.e. positive or negative) and the retention time at which each compound is eluted through the chromatographic column were reported too.

5.2.6 Reference material and quality control

A reference material for each fruit part (e.g. receptacle and drupelet) and for each fruit phenotype (e.g. artificial 'crumbly' and control) was generated by pooling together in equal amounts the two different development stages (e.g. green fruit and red fruit). In total 12 reference sample materials, for each sample (i.e. drupelet control, drupelet 'crumbly', receptacle control and receptacle 'crumbly') 3 replicates were produced. Each reference sample weighed 50 mg. It was processed (i.e. hormones extraction and sample

preparation) exactly as for the experiment samples and stored dried at -80 °C, ready to be reconstituted in 40 µL (30% acetonitrile) the same day of the analysis.

The accuracy and precision of the analytical method was assessed with quality control (QC) sample which consisted of identical aliquots (10 µL) of each of the 12 reference samples pooled together. The QC sample was repeatedly (10 times in total) injected through the respective sequence at the beginning, before and after the reference material samples as well as before and after the experiment samples.

5.2.7 Calibration curves and linearity

The calibration curves were run in triplicate, in receptacles matrix (RM) and in drupelets matrix (DM) both green and red stage combined in equal part and in pure solvent (blank (B) 50% methanol v/v). The curves were obtained by 12 serial dilutions of the standards mix solution, containing all the eighteen target phytohormones and the nine isotopically labelled internal standards. In Figure 5.2 was reported a diagram to describe the procedure adopted to prepare the calibration curves samples for both blank solvent and matrices (i.e. drupelet and receptacle.

The most important aspect was keeping the concentration of the labelled internal standards (ISTD) constant while serially halving the concentration of the 18 standard target compounds. The approach was to use as eluent for the dilution a solution with the same concentration of ISTD, in pure solvent and in matrix, helped to achieve such goal.

Data were first processed in MassHunter software (Agilent, USA) to correct manually, potential errors in peak area integration and only after that, they were exported to a Microsoft Excel spreadsheet for calculating the regression parameters and produce the curve plot for each of the 18 compounds of interest. The peak area of each target compound was normalized with the peak area of the corresponding labelled internal standard; not all the target compounds had its exact labelled counterpart but there was at least one for each of the six classes of plant hormones (i.e. abscisic acid, auxins, jasmonates, gibberellins, salicylates and cytokinin). The intervals that best fitted the linear regression model were chosen and, for those compounds with better ionising properties and subsequent better detection a broader linear dynamic range was found.

The calibration curves were calculated in both solvent (50% methanol) referred to as (blank) and in receptacle and drupelets matrix (green plus red stage fruit), referred to as (matrix). Two extra calibration curves were calculated for each target compound and the idea behind this was to produce curves where the effect of the matrix was removed. This was achieved in two different ways referred to as (matrix – sample) in which the peak area of the target compound, for each of the 12 serial dilutions, was first subtracted from the average peak area in the analysed sample and the difference obtained was divided by the peak area of the labelled internal standards from the samples of the calibration curves run in matrix. The second way referred to as (matrix – samples relative ratio) was calculated as the difference between the normalized peak area of the target compound in the calibration curves samples run in matrix and the averaged peak area (normalized) of the target compounds in the samples analysed. The type of calibration curve that showed the broader linear concentration range and the best correlation coefficient was used for the detection/quantification of the target compounds in the samples analysed.

<table>
<tr><th colspan="4">Blank</th><th colspan="4">Matrix (drupelet or receptacle)</th></tr>
<tr><td>standard mix (doubled [STD[a]])</td><td>A</td><td>internal standard mix (doubled [ISTD[b]])</td><td>B</td><td rowspan="2">one sample[c] spiked with ISTD[b] and reconstitute with 40 µL of STD[a] mix solution</td><td rowspan="2">D</td><td rowspan="2">dilution eluent: sample[c] spiked with ISTD[b], processed and then reconstituted with 50% CH_3OH</td><td rowspan="2">E</td></tr>
<tr><td>volume</td><td>0.1 mL</td><td>volume</td><td>0.3 mL</td></tr>
<tr><td colspan="4">dilution eluent</td><td>volume</td><td>40 µL</td><td>volume</td><td>0.28 mL</td></tr>
<tr><td colspan="3">0.260 mL (B) + 0.260 mL 50% CH_3OH</td><td>C</td><td colspan="4"></td></tr>
<tr><td colspan="3">volume</td><td>0.52 mL</td><td colspan="4"></td></tr>
<tr><td>serial dilutions</td><td colspan="3"></td><td>serial dilutions</td><td colspan="3"></td></tr>
<tr><td>12</td><td colspan="3">40 µL (A) + 40 µL (B)</td><td>12</td><td colspan="3">40 µL (D)</td></tr>
<tr><td>11</td><td colspan="3">40 µL (dilution 12) + 40 µL (C)</td><td>11</td><td colspan="3">20 µL (dilution 12) + 20 µL (E)</td></tr>
<tr><td>-</td><td colspan="3">-</td><td>-</td><td colspan="3">-</td></tr>
<tr><td>1</td><td colspan="3">40 µL (dilution 2) + 40 µL (C)</td><td>1</td><td colspan="3">20 µL (dilution 2) + 20 µL (E)</td></tr>
<tr><td colspan="8">[a]standards of target compounds – [b]isotopically labelled standards of target componds
[c]50 mg of plant material (i.e. drupelet or receptacle processed in the same way of the samples prepared for the analysis)</td></tr>
</table>

Figure 5.2: Schematic diagram of the procedure used for the preparation of the samples for the calibration curves in both bank and matrix.

The goal was to keep constant the concentration of the isotopically labelled internal standards (ISTDs) while serially halving the concentration of the 18 target compounds. For the samples in blank solvent, two solutions, named (**A)** and (**B)**, containing double the expected concentration of target compounds and ISTDs, were prepared. The highest serial dilution was prepared by mixing the same volume (40 µL) of the solutions **A** and **B.** In parallel, the diluent solution (**C**) was prepared by mixing same volumes of the solution **B** and a solution of 50% of methanol. The new solution **C** (diluent) would have always the same concentration of ISTDs of the solution to be diluted allowing to keep constant the ISTDs concentration while, on the other hand, halving serially that of the target compounds. For the calibration samples run in matrix the procedure was a little bit different. Eight samples, each weighing 50 mg, (mix of green and red receptacle or drupelets) were spiked with ISTDs in exact the same concentration required for the analysis. This calibration samples were processed in the same way of the experiment samples. After processing, one sample was reconstituted with 40 µL of solution containing the 18 target compounds, forming the solution (**D**). The solution **D** will be the first serial dilution of the calibration curves, the one with highest concentration. The other 7 calibration matrix samples were reconstituted, each, with 40 µL of 50% methanol and then pooled together to form the solution (**E**); the diluent of the serial dilution. The solution **E** having always, during the dilution**,** the same concentration of ISTDs of the serial dilutions, helped keeping constant the ISTDs concentration while that of the target compounds halved serially.

5.2.8 Limit of detection and limit of quantification

Terms such, Limit of Detection (LoD) and the Limit of Quantification (LoQ) were used to indicate the smallest concentration of a target compound that could be reliably measured by an analytical procedure (Armbruster and Pry, 2008). In practice, an assay is not able to measure the analyte concentration down to zero, there must always be a sufficient analyte concentration that produces a signal that can be reliably distinguished from the 'noise' that is the signal produced in the absence of the analyte (Armbruster and Pry, 2008). In the case of linear calibration curves, it can be assumed that the instrument response (y) and the analyte concentration (x) are linearly related and that this relation can be expressed by the equation (Shrivastava and Gupta, 2011):

$$y = a + bx$$

The slope of the curve (b) which represents the sensitivity of the instrument can be used to calculate LoD and LoQ (Shrivastava and Gupta, 2011).

LoD was then expressed as $LoD = 3S_a/b$ while $LoQ = 10S_a/b$, where S_a is the standard deviation of the instrument response and it can be estimated by the standard deviation of the y-residuals of regression lines (Shrivastava and Gupta, 2011).

For quadratic relation (i.e. $y = ax^2 + bx + c$) a similar approach was used. The gradient or slope of a quadratic function is not constant. It is possible to calculate the gradient of a curve (at a point) by calculating the gradient of the tangent to the curve at that specific point. For ($y = ax^2 + bx + c$) the gradient function was calculated by differentiation ($d_y/d_x = 2ax + b$). Once the gradient function was calculated, the lowest point of the quadratic function was chosen to calculate the gradient of the curve at that point. The calculated gradient or slope was then used to estimate LoD and LoQ accordingly to the formulas $LoD = 3S_a/b$ while $LoQ = 10S_a/b$, where S_a is the standard deviation of the instrument response and it can be estimated by the standard deviation of the y-residuals of regression lines. S_a was calculated using the polynomial regression function in GenStat (VSN International, UK).

5.2.9 Statistical analysis

In order to assess whether or not the two different kind of treatments (i.e. damage of the receptacle and control) were significantly different, and because the artificially induced crumbly fruits were the number of fruits that crumbled out of those treated, a generalised linear model with a binomial distribution was performed (GenStat, VSN International, UK). All the statistical analysis of the metabolomic measurements (i.e. ANOVA, principal component analysis, linear regression, were run in GenStat (VSN International, UK). The analysis of variance, of the 120 samples (see Figure 5.3 below for further details about the experiment design), whose phytohormones concentration were measured with the analytical method described in section 5.2.5 of this chapter, was performed, assuming a completely randomised experiment design, with a blocking structure (i.e. plant + batch) and with treatment*stage*part interaction.

The polynomial regression analysis was performed to calculate the equation of calibration curves required for the detection/quantification of the amount of target phytohormones in the tested samples. A linear regression with quadratic polynomial model was used setting the instrument response (i.e. ratio between analyte peak area and its corresponding internal standard) as the independent variable while the analyte concentration was set as the dependent variable.

The PCA analysis was based on correlations which mean that each variate was standardized. by subtracting its mean and then dividing such difference divided by its standard deviation); such an approach is used when the variates do not share the same scale and show very different amount of variation, GenStat manual (VSN International, UK). The sixty samples per stage (i.e. green and red fruit) were analysed together so that two principal component analysis were performed, one for the drupelet and receptacle of the green stage fruit and another one for the same sample parts but of the red stage. For any PCA performed, the five detected/quantified phytohormones were the variates, while the 2 parts per stage were the samples.

A generalised linear model with binomial distribution and logit link function was performed in GenStat (VSN International, UK) to verify that the treatment (damage of the receptacle) could produce, confidently, artificial ‘crumbly’ fruit compared to the (control) treatment.

5 plants	**X**	**two treatments**		**X**	**two parts**		**X**	**two stages**		**X**	**three replicates**			**total samples**
		‘crumbly’	control		receptacle	drupelet		green	red		a	b	c	120

Figure 5.3: Schematic diagram of the design of the ‘crumbly’ fruit induction experiments.

Five plants (cultivar Glen Ample) were grown in controlled environment. Flowers were emasculated and half were pinned with a needle to damage the receptacle and then hand pollinated with brush while the other half were simply hand pollinated to form the control samples. Three fruits at both green and red stage of both receptacles and drupelets were collected from each plant. In total 120 samples were produced and analysed.

5.3 Results

5.3.1 Crumbly fruit induction experiments

Experiments were conducted in a controlled environment (section 5.2.1) and in total 224 flowers were emasculated, damaged by pinning the tip and side of the receptacle with a needle and then hand pollinated with a small painting brush while the others were simply hand pollinated to form the control samples (Table 5.7). Of the six plants (Glen Ample cultivar) used in the experiment, one did not grow properly, producing only a few laterals on the top of the cane while all the other plants produced enough flowers to allow an adequate number of replicates to be used for the analysis. A general linear model with binomial distribution and logit link function was performed to verify that the artificial induced 'crumbly' fruit were the effect of the damage employed rather than a random event. Each response (i.e. crumbly, normal and no fruit), for details see Table 5.6 below, was analysed relative to the number of treated flower buds. The dispersion was estimated from the data and for each response there was a significant effect of treatment; the predictions were made fitting a model with the treatment and the proportions for each predicted response were reported in Table 5.6 below.

Treatment	crumbly		normal		no fruit	
	prediction	S.E.	prediction	S.E.	prediction	S.E.
control	0.0347	0.01834	0.8505	0.03448	0.1121	0.03051
[a]R damage	0.8803	0.02999	0.0000	0.00006	0.1197	0.03

[a]damage of the receptacle with a pin

Table 5.6: Prediction from regression model of the three possible responses (i.e. 'crumbly' no fruit and normal fruit) from the 'crumbly' induction experiments.

Each response (i.e. 'crumbly', normal fruit and no fruit) to the designed treatment (i.e. damage of the receptacle and control) was analysed with a generalised linear model with a binomial distribution and logit link function. The analysis was performed to verify that the treatment (damage of the receptacle) could produce, confidently, artificial 'crumbly' fruit compared to the (control) treatment. With respect to the 'crumbly' and normal responses, the treatment (damage of the receptacle) was significant with $p<0.001$. There was no significant difference between the control and treated plants for the numbers failing to produce fruit.

Focusing on the artificial 'crumbly' fruit, the damage of the receptacle was significantly different from the control treatment with $p<0.001$ so the analysis gave confidence that the experiments succeeded in producing artificial 'crumbly' like misshapen fruits.

Plant	treatment	[a]n. treat.s	'crumbly'	no fruit	normal
1	[ł]R	25	22	3	0
	[§]C	21	2	4	15
2	[ł]R	25	24	1	0
	[§]C	21	0	2	19
3	[ł]R	7	6	1	0
	[§]C	6	0	0	6
4	[ł]R	27	22	5	0
	[§]C	26	1	3	22
5	[ł]R	6	6	0	0
	[§]C	7	0	0	7
6	[ł]R	27	23	4	0
	[§]C	26	1	3	22
total		224	107	26	91

[a]number of flowers treated
[ł]damage of the receptacle
[§]control

Table 5.7: 'Crumbly' fruit induction experiment results.

The experiments were run in a controlled environment (growth room) on six Glen Ample plants randomly positioned in the cabinet. Plants were kept in ideal growing conditions and two different treatment applied, **R** (damage of the receptacle) and **C** (control).

5.3.2 Phytohormones detected/quantified in crumbly induced samples

The analytical method described in the previous section (5.2 Materials and methods) of this chapter, was developed to simultaneously analyse 18 different plant hormones but of these only five (i.e. abscisic acid, gibberellin A_1, indole-acetic acid, salicylic acid and gibberellin A_4) were detected and were only partially quantified, resulting in a semi-quantitative analysis of the hormone levels in the samples tested. The instrument response

to all the remaining 13 target compounds was indistinguishable from the instrument noise and consequently, their analysis was not progressed further.

For technical reasons all 120 samples could not be analysed at the same time, (analysis lasting29 minutes per sample requiring about sixty hours (approximately 3 days) for the 120 samples. The analysis design also required quality control (QC) and reference sample material to be analysed before, in between and after the experimental samples, the reason being that as they were composed of exactly the same material, they would have helped in monitoring instrumental behaviour thereby facilitating corrections of the instrumental analytical response. In total by considering the 12 serial dilution to be run in blank (pure solvent), receptacle matrix and drupelet matrix, totalling 36 samples, plus the 30 QC, the 108 references and the 120 experiment samples, the final minimum amount of sample to be analysed by the triple quadrupole LC-MS instrument would have been almost 300 meaning more than six days in continue operation mode; impractical for the instrument. It was therefore decided to run the samples in three different batches, each corresponding to the three biological replicates available for the measurements.

All the three classes of samples, artificially 'crumbly' induced experiment samples, quality control and reference material samples were subjected to analysis of variance (ANOVA) in order to identify and then remove potential outliers. Once those were removed, scatter plots were run for each of the five target compounds to verify that a stable instrument response was occurring along the whole duration of the analysis and so exclude any evident effect of the sequence order on the analysis. No difference in the instrument performance was found along the whole operating time for all the three batches of samples analysed. No evident alteration in the instrument response related to the samples sequence order was found for all the five target compounds detected/quantified. In Figure 5.4, the scatter plot related to salicylic acid is identified as an example. In the plot, the quality control samples were reported as green crosses while the reference materials samples as a sky-blue squares (drupelets) and as a pink rhombuses (receptacle). Except for a few samples, both drupelet and receptacle of reference materials, all the other samples (QC and reference) were well horizontally aligned suggesting that no evident instrument response alteration, in respect to the sequence order of the analysis, occurred during the analysis. This approach, relatively fast and simple suggested that the variability in the data was biological and not technical, apart from the instrument limitations, and so no data adjustment was necessary for all the samples and for all the five different target compounds.

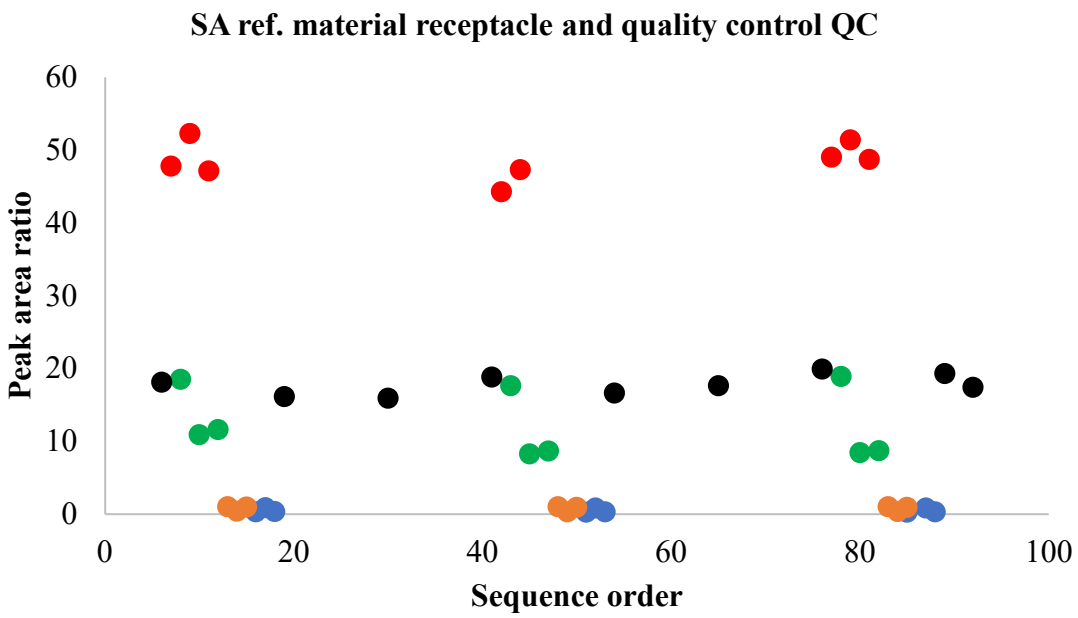


Figure 5.4: Scatter plot (batch 3) of the whole data set for the target compound salicylic acid (SA). In the plot, the quality control (QC) samples were reported as black circles, the receptacle reference (REF) materials samples as red and green circles respectively for control and artificial 'crumbly' samples while the drupelet reference (REF) as orange and blue circles respectively for the control and artificial 'crumbly' samples. On the horizontal axis were plot the sequence order for the 100 analysis and on the vertical axis the instrument response for the compound under analysis (SA in this example). This allowed to visually compare the instrument responses along the duration of the analyses. QC samples and reference material samples were specifically designed to detect potential alteration in the instrument functioning; they were the same samples re-analysed at different times, before, between and after the experiment samples. The plot clearly showed how the black circles (QC) as well as the various REF material samples, except few samples, were well horizontally aligned indicating that the instrument response did not change significantly during the time and that the measurement at the beginning and those at the end did not require any adjustment.

The metabolomic data (i.e. peak areas ration between analyte and corresponding ISTD) related to the 120 crumbly induction experiment samples, with outliers removed , were subjected to ANOVA with a blocking structure (i.e. plant + batch) and with treatment*stage*part interaction design. This was done to aid identification of differences in the target phytohormones, not only at phenotype (i.e. artificial crumbly and normal fruit) level, but at the interaction of all the three factors (i.e. type, stage and part). The goal was not just to find any significant difference in the hormones level, between the two phenotypes (i.e. artificial crumbly and normal fruit), but also to determine whether these differences would be found at an early or late stage of the fruit development as well as in which part (i.e. receptacle and drupelet). This was to get as much information as possible on the hypothesis around the hormone based regulation of the development process, proposed in the introduction of this work and to which the crumbly induction

experiments were designed. According to this proposed fruit development model, the growth of fertilised ovules would be coordinated by the receptacle that acting as leading hub, which would synchronise and regulate the growth of all the fertilised ovaries.

Abscisic acid (ABA) and Gibberellin acid A_1 (GA_1) were the only target plant hormones whose differences were statistically significant at about 95% of confidence level (see Table 5.8), respectively in receptacle green stage and in drupelet red stage; with ABA being significantly higher in the green receptacle of the artificially induced 'crumbly' samples (see Figure 5.5) while GA_1 being significant lower in the red drupelets of the artificially induced 'crumbly' samples (see Figure 5.6).

Such findings gave confidence that ABA and GA1 might play an important role in the process leading to the formation of misshapen 'crumbly' like fruits; with ABA being involved at early stage and in the receptacle while GA1 at a later stage but in the drupelets.

Compounds	treatment	stage	part		[a]d.f.	[b]rep.	[c]S.E.
			drupelet	receptacle			
ABA	crumbly	green	0.905	1.697*	1	15	0.0518
	control		0.849	1.371*			
	crumbly	red	2.101	2.401			
	control		2.100	2.393			
GA_1	crumbly	green	0.013*	0.594	1	15	0.086
	control		-0.096*	0.560			
	crumbly	red	-0.178	0.619			
	control		0.067	0.641			
IAA	crumbly	green	0.504	-0.058	1	15	0.0998
	control		0.641	0.077			
	crumbly	red	0.408	-0.387			
	control		0.324	-0.407			
SA	crumbly	green	-0.235	1.591	1	15	0.0779
	control		-0.196	1.645			
	crumbly	red	-0.478	1.116			
	control		-0.417	1.262			
GA_4	crumbly	green	-1.205	-0.624	1	15	0.118
	control		-1.231	-0.526			
	crumbly	red	-1.220	-0.632			
	control		-1.253	-0.444			

*differences being statistically significant at 95% of confidence levels
[a]degree of freedom
[b]number of replicates
[c]standard error of differences of means from ANOVA interaction (treatment*stage*part) effect

Table 5.8: Analysis of the variance for the five target phytohormones detected/quantified with the LC-MS analytical method.

ANOVA table of means for the treatment*stage*part interaction effect of the 120 crumbly induction experiment samples related to the five phytohormones detected/quantified with the analytical method developed during this work. ABA and GA_1 were the only target phytohormones showing significant difference, between the two phenotypes (i.e. artificial crumbly and control) at about 95% of confidence level, respectively in green receptable (ABA) and red drupelet (GA_1).

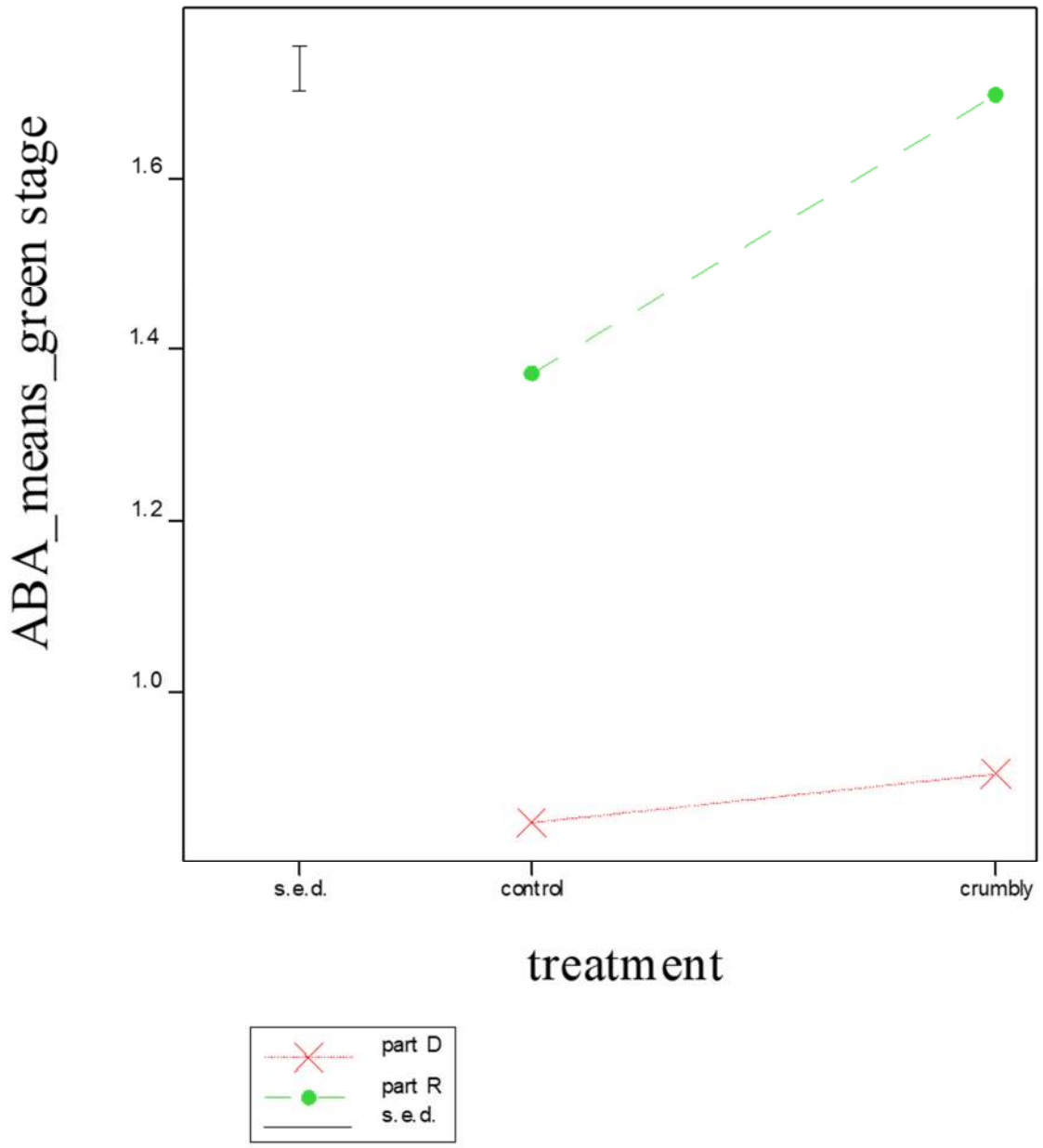


Figure 5.5: : Means line plot of ABA peak areas for receptacle and drupelets at green stage in both control and artificially induced 'crumbly' samples.

ABA means for treatments (i.e. damage of receptacle 'crumbly' and normal fruit 'control') for the two different parts (i.e. drupelet and receptacle) at green stage fruit. In the receptacle samples (green dashed line), the averaged peak area, of the artificially induced 'crumbly' samples, was significatively higher at about 95% of confidence levels compared to that of control samples (normal shape fruits).

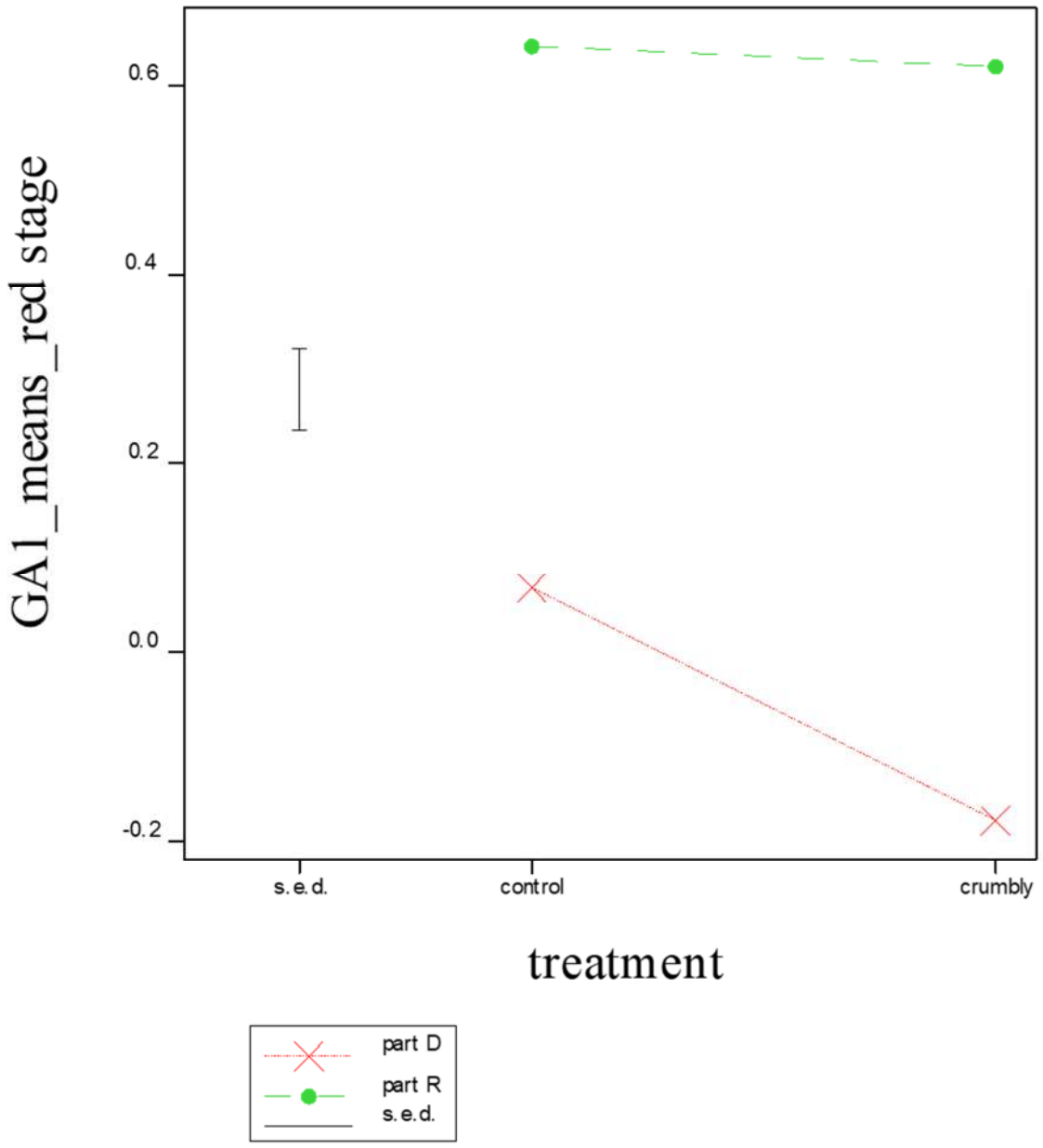


Figure 5.6: Means line plot of GA_1 peak areas for receptacle and drupelets at red stage in both control and artificially induced 'crumbly' samples.

GA_1 means for treatments (i.e. damage of receptacle 'crumbly' and normal fruit 'control') for the two different parts (i.e. drupelet and receptacle) at red stage fruit. In the drupelet samples (red dotted line), the averaged peak area, of the artificially induced 'crumbly' samples, was significative lower, at about 95% of confidence levels, compared to that of control samples (normal shape fruits).

The analysis of the variance highlighted useful information about the potential hormonal regulation of the normal fruit development and in fact although, with respect to phenotype (i.e. artificially induced 'crumbly' and normal fruit), no further significant differences were found for the other compounds; the analysis still showed significant differences found at the level of stage*part interaction (Table 5.9).

The hypothetic model for fruit development, described in the introduction (see section 5.1 of this chapter) proposed the receptacle as a coordinator, a central hub, which by means of molecular crosstalk with the fertilised ovaries, mediated by specific signal (hormones), led the simultaneous growth of all the drupelet. In this manner differences in the fertilization time would not affect the development process avoiding the creation of a gap in the growing process between ovaries fertilized at different times that might cause the formation of misshapen fruits.

Bar charts of the averaged peak areas for GA_1, IAA and GA_4 (Figure 5.6) and for ABA and SA (Figure 5.7) gave a clear and immediate overview of the results of the phytohormones analysis for the five detected/quantified target compounds. In Table 5.9 the predicted means of the analysis of variance for the stage (i.e. green and red fruit) per part (i.e. drupelet and receptacle) were reported. The data from the metabolomic analysis clearly showed that in the receptacle, the hypothetical fruit growth regulator, the phytohormones that were present in higher amounts compared to the drupelet were ABA, SA and the two gibberellins (i.e. GA_1 and GA_4), with these last two always being present at the same amount regardless of the stage. While ABA and SA were higher in red and green fruit, respectively. IAA was the only phytohormone showing higher concentration in drupelet, especially at green stage. In Table 5.11 a schematic summary of the analysis results for the five target phytohormones was reported.

Such findings would suggest that in a hypothetical fruit development model, SA and IAA might play a crucial role at the beginning of the fruit development with the first acting in the receptacle and the second in the drupelets. The two gibberellins being neutral since their concentration was not significantly different between the two stages, and ABA playing a crucial role in the last stage of fruit development. This is not unexpected, since in non-climacteric fruit, such as raspberry, ABA and not ethylene, plays the main role in fruit ripening.

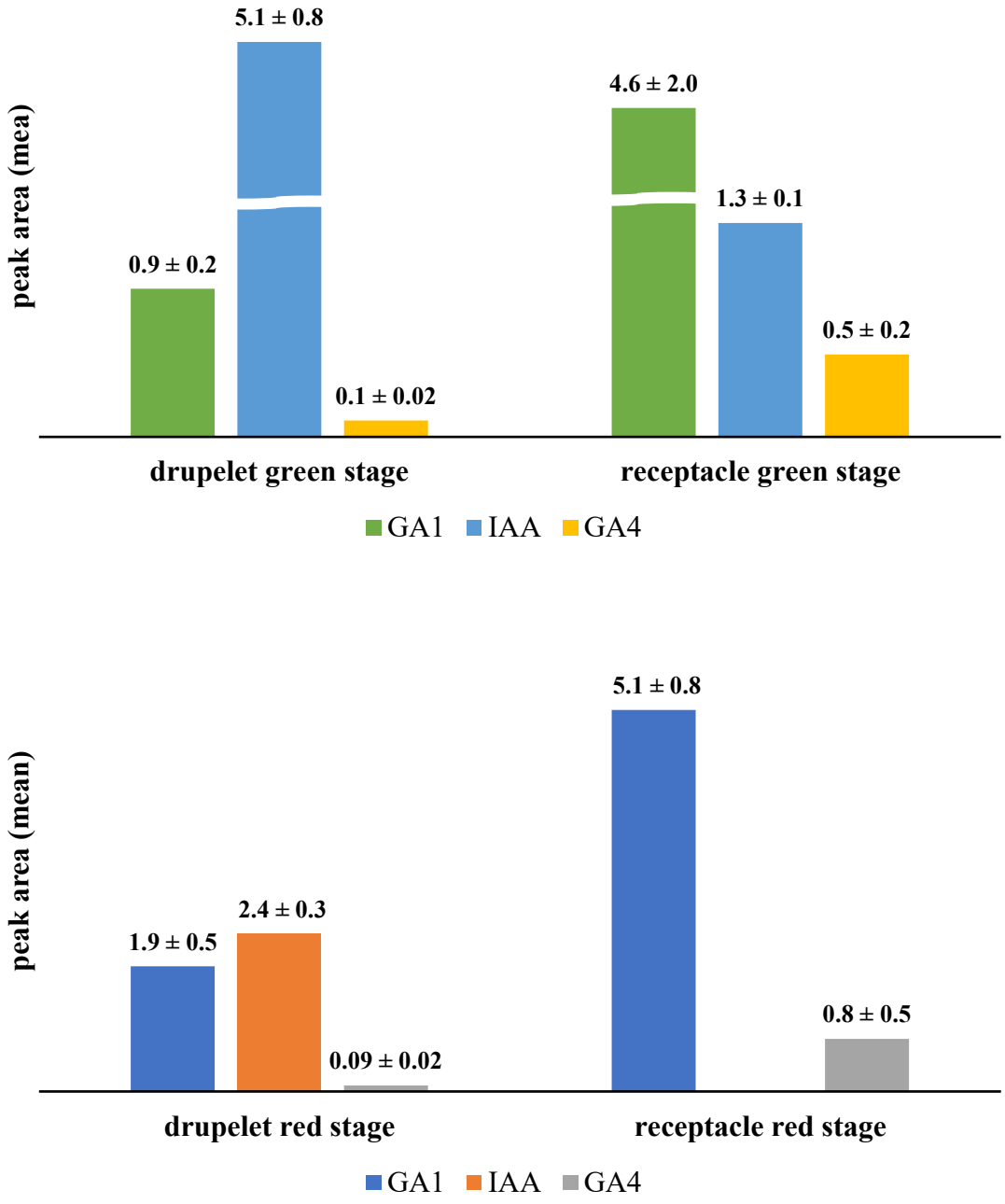


Figure 5.7: Bar chart of GA1, IAA and GA4 at both green (top) and red (bottom) stage fruit.

Bar chart of the averaged peak area ratio (i.e. analyte peak area/corresponding ISTD peak area) for three out of the five detected/quantified phytohormones. The differences between the two parts (i.e. drupelet and receptacle) were always significative with gibberellins peak areas ratio always higher in receptacle while IAA was higher in drupelet. For what concerned the differences at the stage (i.e. green and red fruit) per part (i.e. drupelet and receptacle) interactions, while gibberellins were not significative different between green and red stage, IAA was significative higher at green stage.

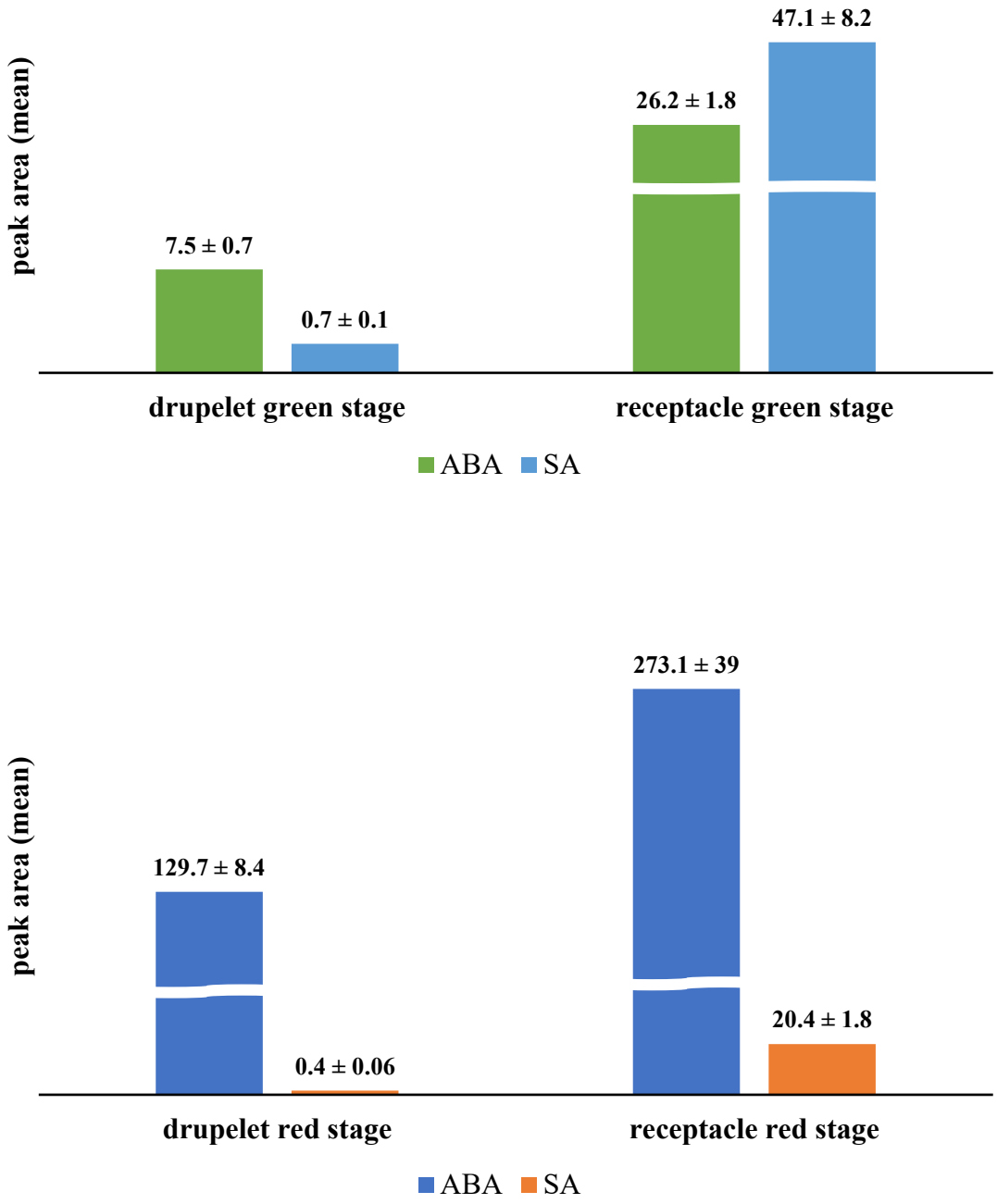


Figure 5.8: Bar chart of ABA and SA in drupelet and receptacle samples at both green (top) and red (bottom) fruit stages.

Bar chart of the averaged peak area ratio (i.e. analyte peak area/corresponding ISTD peak area) for two (i.e. ABA and SA) out of the five detected/quantified phytohormones. The differences between the two parts (i.e. drupelet and receptacle) were always significative with the peak areas of both the two phytohormones bigger in receptacle. For what concerned the differences at the stage (i.e. green and red fruit) per part (i.e. drupelet and receptacle) interactions, both ABA and SA were significative different between green and red stage, with ABA peak areas ratios hugely higher at red stage while SA peak area, on the contrary, was bigger at green stage.

Compounds	stage	part		[a]d.f.	[b]rep.	[c]S.E. (stage*part)
		drupelet	receptacle			
ABA	green	0.877*	1.534*	1	30	0.0367
	red	2.100*	2.397*			
GA_1	green	-0.042	0.577	1	30	0.0608
	red	-0.055	0.630			
IAA	green	0.572*	0.01*	1	30	0.0706
	red	0.366*	-0.397*			
SA	green	-0.215*	1.618*	1	30	0.0551
	red	-0.448*	1.189*			
GA_4	green	-1.218	-0.575	1	30	0.0834
	red	-1.237	-0.538			

*differences being statistically significant at 95% of confidence levels
[a]degree of freedom
[b]number of replicates
[c]standard error of differences of means from ANOVA interaction (stage*part) effect

Table 5.9: ANOVA table of means for the stage*part interaction of the five detected/quantified target phytohormones.

For each target phytohormones (i.e. ABA, GA_1, IAA, GA_4 and SA) was reported, together with the corresponding standard error (S.E.), the predicted means from the ANOVA analysis for the stage (i.e. green and red fruit) per part (i.e. drupelet and receptacle) interaction for all the 120 samples. The analysis showed that the differences in the predicted means were significant at 95% confidence levels for three out of the five compounds (i.e. ABA, IAA and SA).

Compound	receptacle	drupelet	green	red
ABA	high	low	low	high
GA_1	high	low	=	=
IAA	low	high	high	low
SA	high	low	high	low
GA_4	high	low	=	=

Table 5.10: Schematic diagram of the differences in the level of the five detected/quantified target phytohormones.

For each of the five detected/quantified target phytohormones, differences being significant at 95% confidence level, were reported for both parts (i.e. drupelet and receptacle) and stages (green and red fruit). ABA receptacle > drupelet with red fruit > green fruit, GA_1 receptacle > drupelet with no significant differences were found in stages, IAA drupelet > receptacle with green fruit > red fruit, SA receptacle > drupelet with green fruit > red fruit and finally GA_4 receptacle > drupelet with no significant differences found in stages.

5.3.3 Principal component analysis green stage normal shape fruit

The five phytohormones (i.e. ABA, GA_1, IAA, SA and GA_4) detected/quantified with the metabolomic analysis were subjected to multivariate analysis (Principal Component Analysis) in order to try to identify those compounds potentially playing a major role in the developmental process of normal fruit. Always referring to the proposed model of fruit growth, coordinated and regulated by the receptacle, the PCA analysis was performed for both the two stages (i.e. green and red fruit), analysing together the two parts (i.e. drupelet and receptacle) with the aim of finding those primarily involved in the regulating process and the site (i.e. drupelet or receptacle) of their action; to try to better understand by means of which hormones the receptacle eventually would regulate and coordinate the growth of all the fertilised ovaries together.

The PCA analysis was run in GenStat (VSN International, UK) with the five hormones being the variates while the three replicates of the five plants at the two different stages (i.e. green drupelet and green receptacle) were the samples. The analysis was performed using the procedure described in section two of this chapter). The result of the analysis showed that the first two principal components explained respectively about the 56% and the 21% of the variation (77% of the total variation) and to evaluate the hormones responsible for this separation, direction and magnitude of the vectors in the PCA were interpreted. Focusing on the first principal component (PC-1) that determines the separation of the sample along the horizontal axis, in the scatter plot having the PC-1 on the x-axis, all the drupelets samples were displayed on the left of the diagram and clearly separated from the samples related to the receptacle that stood on the right side (see Figure 5.8).

The latent vector loadings were: +0.50535 x GA_1 +0.50088 x SA -0.49196 x IAA +0.47081 x ABA +0.17336 GA_4, the result of the analysis showed that apart from GA_4, all the other four hormones seemed to be involved at the same level since the magnitude of their vectors are very similar, with the only exception being IAA whose direction is negative meaning that it is negative correlated to the other vectors. This finding is in line with the result from the ANOVA analysis that gave highlighted the significantly higher content of this hormone in the drupelet while the others were more abundant in the receptacle. Essentially the PCA showed that the crosstalk between GA_1, SA, IAA and ABA, at green stage fruit, were all very important for the processes regulating the growth and development of raspberry fruits in both the two analysed parts (i.e. drupelet and

receptacle); with the difference that IAA would contribute more to the growth of the drupelets while the other three more in that of the receptacle.

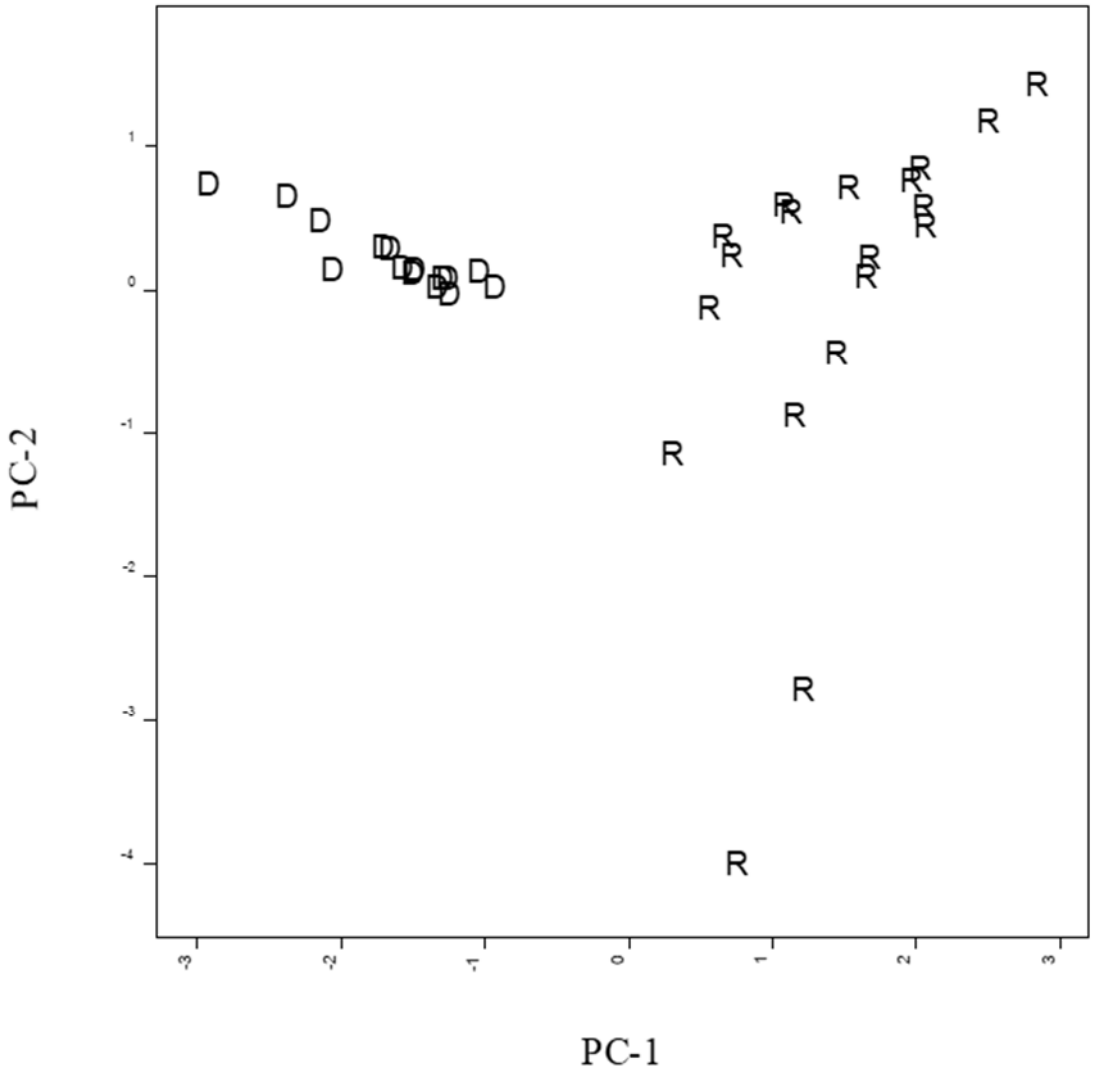


Figure 5.9: Principal component analysis (PCA) scatter plot of drupelet and receptacle samples at green stage for the five detected/quantified target phytohormones.

Scatter Plot showed the complete and clear separation along the first principal components for the two different fruit parts (i.e. drupelet (**D**) and receptacle (**R**) at green stage), with the drupelets on the left and the receptacles on the right of the plot.

5.3.4 Principal component analysis red stage normal shape fruit

PCA was also performed on the samples (i.e. drupelet and receptacle) from the late stage (red fruit), so again the five phytohormones (i.e. ABA, GA_1, IAA, SA and GA_4), detected/quantified with the LC-MS analytical method. The phytohormones were the variates while the three replicates from the five plants for both the two different stages (i.e. green and red fruit) were the samples. The result of the analysis showed that the two first principal components explained respectively about 47% and 23% of the variation (70% of the total), the evaluation of the hormones mainly responsible for this separation was achieved by looking at direction and magnitude of the vectors in the PCA.

Focusing on the first principal component, the one that separated completely the two groups of samples (i.e. drupelet and receptacle), see Figure 5.9, the latent vector loadings were: +0.56784 x SA +0.47952 x ABA +0.45322 x GA_4 -0.38514 x IAA +0.30639 x GA_1. This time, the results of the analysis showed that except for GA_1 that seemed to have a smaller impact on late stage of fruit development, it was SA that proved to have the greater effect, the magnitude of its vector in the analysis was by far the highest but ABA, GA_4 and IAA too were involved with IAA having opposite direction indicating its negative correlation with the other three hormones. This result was not surprising since as the ANOVA analysis indicated, IAA was more abundant in the drupelet while all the other plant hormones were more abundant in the receptacle (Table 5.11). In synthesis, at red stage, the interplay between SA, ABA, GA_4 and IAA regulates the processes leading to fruit growth and development in both the two fruit parts (i.e. drupelet and receptacle). As for the green stage even at this latest stage IAA contributed more to the growth of the drupelet while the other phytohormones more to that of the receptacle but despite that, all these five phytohormones were involved in the growth of both receptacle and drupelets.

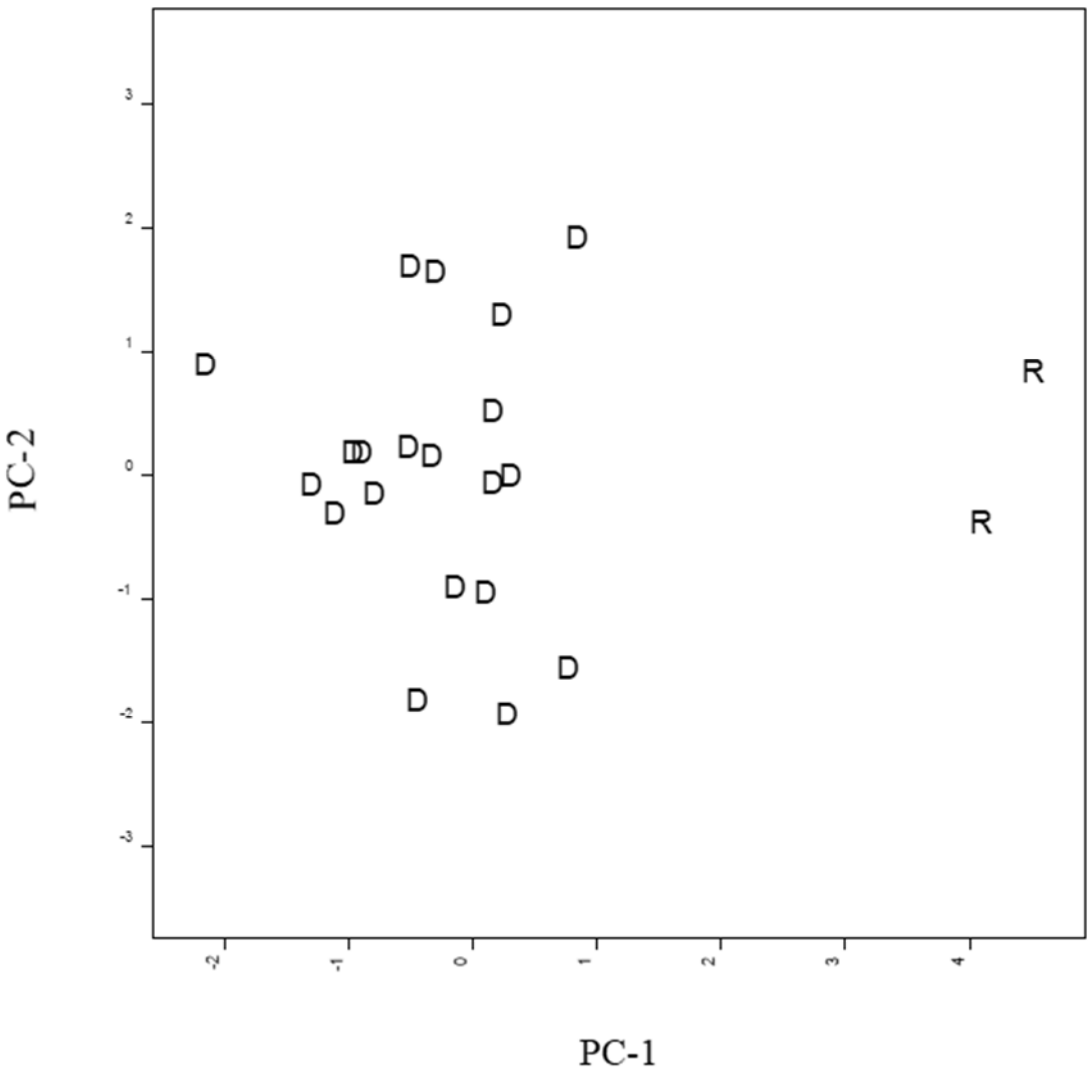


Figure 5.10: Principal component analysis (PCA) scatter plot of drupelet and receptacle samples at red stage for the five detected/quantified target phytohormones.

Scatter Plot showed the complete and clear separation along the first principal components for the two different fruit parts (i.e. drupelet (**D**) and receptacle (**R**) at red stage), with the drupelets on the left and the receptacles on the right of the plot.

5.3.5 Phytohormones semi-quantification

Targeted LC-MS analysis proved to be successful only for five out of the eighteen target plant phytohormones for which the analytical method was designed and developed. Calibration models were produced to help derive concentration values from the instrument response (i.e. ratio peak areas analyte/corresponding ISTD). When those predicted values were higher than the limit of detection, an estimation of the amount for the specific analyte was performed.

DRUPELET						
Compound	**equation**	**R^2**	**[a]LOD**	**[b]LOQ**	**linear range**	**batch**
ABA	$y = -4E\text{-}06x^2 + 0.008x + 0.0763^{§}$	0.9998	0.09625	0.14262	0.09 – 1.5	1
GA_1	$y = 0.0002x^2 + 0.0058x – 5E\text{-}05^{ł}$	0.9995	0.00214	0.00745	0.00097-0.06	1
IAA	$y = 2E\text{-}05x^2 + 0.0009x + 0.0002^{ł}$	0.9998	0.00046	0.00119	0.0002-0.02	1
SA	$y = 0.0213x^2 + 0.2876x + 0.0002^{§}$	0.9994	0.01857	0.06209	0.005-0.6	1
GA_4	$y = -0.2243x^2 + 0.9831x + 0.0054^{§}$	0.9999	0.01713	0.04424	0.03-0.5	1
ABA	$y = 2E\text{-}06x^2 + 0.0029x - 0.0039^{ł}$	0.9996	0.0144	0.0576	0.003-0.75	2
GA_1	$y = 0.0053x^2 + 0.0782x – 0.0004^{ł}$	0.9991	0.00853	0.02999	0.0009-0.25	2
IAA	$y = 2E\text{-}05x^2 + 0.0049x - 0.0004^{ł}$	0.9996	0.00129	0.0053	0.00097-0.625	2
SA	$y = -0.0009x^2 + 0.171x + 0.0015^{§}$	0.9984	0.03749	0.12114	0.019-0.625	2
GA_4	$y = 0.103x^2 + 0.3624x + 0.0034^{§}$	0.9995	0.01902	0.05687	0.014-0.47	2
ABA_high	$y = 3E\text{-}06x^2 + 0.0094x - 0.0262^{§}$	0.9876	0.19524	0.72386	0.02-1.5	3
ABA_low	$y = 7E\text{-}06x^2 + 0.0029x + 0.0014^{ł}$	0.9998	0.00481	0.01286	0.003-0.2	3
GA_1	$y = -0.0013x^2 + 0.0335x – 0.001^{ł}$	0.9997	0.00210	0.00920	0.0019-0.125	3
IAA	$y = -2E\text{-}05x^2 + 0.0047x – 1E\text{-}05^{ł}$	0.9996	0.00142	0.00473	0.0002-0.06	3
SA	$y = -0.0019x^2 + 0.1671x - 0.0043^{ł}$	0.999	0.02356	0.08817	0.02-0.6	3
GA_4	$y = 0.4877x^2 + 0.6966x + 0.0022^{ł}$	0.9994	0.01008	0.02902	0.007-0.2	3
[a]limit of detection - [b]limit of quantification - [ł]blank - [§]matrix						

Table 5.11: Calibration curves for the samples run in blank solvent and drupelet matrix matched solvent.

Calibration curves equations for the prediction of phytohormones amount in the crumbly induction experiments. Data from samples run both in 50% methanol (blank) and in matrix matched (e.g. 50 mg of raspberry drupelet processed and then resuspended in 50% methanol) and obviously related to the sole phytohormones detectable/quantifiable with the analytical methods optimised for these target compounds in this work. All the curves revealed a broad linear concentration range, spanning over three orders of magnitude and had a correlation coefficient (R^2) of at least 0.9876. In addition, for each calibration curve was reported the limit of detection (LoD) and the limit of quantification (LoQ).

RECEPTACLE						
Compound	**equation**	**R^2**	**[a]LOD**	**[b]LOQ**	**linear range**	**batch**
ABA	$y = -2E\text{-}07x^2 + 0.0049x - 0.0587^{ł}$	0.989	0.118	0.528	0.047 – 1.5	1
GA_1	$y = 5E\text{-}05x^2 + 0.0069x – 0.0006^{ł}$	0.9995	0.00365	0.01381	0.002-0.125	1
IAA	$y = -6E\text{-}05x^2 + 0.0015x - 0.0001^{ł}$	0.9999	0.00143	0.00446	0.0002-0.009	1
SA	$y = 0.0088x^2 + 0.4924x - 0.1179^{ł}$	0.9936	0.177	1.071	0.16-2.5	1
GA_4	$y = 0.0857x^2 + 0.9194x + 0.0015^{ł}$	0.9996	0.05616	0.18602	0.03-1.9	1
ABA	$y = 4E\text{-}06x^2 + 0.0073x + 0.2697^{§}$	0.9997	0.4487	0.8840	0.375-6	2
GA_1	$y = -0.0007x^2 + 0.0941x - 0.0029^{ł}$	0.9995	0.01016	0.04062	0.00097-0.5	2
IAA	$y = 0.0002x^2 + 0.0043x – 9E\text{-}05^{ł}$	0.993	0.00134	0.00488	0.0002-0.02	2
SA	$y = 0.0019x^2 + 0.127x + 0.0925^{ł}$	0.9995	0.46782	1.46035	0.3-10	2
GA_4	$y = -0.0175x^2 + 0.4179x - 0.0371^{§}$	0.9971	0.08575	0.36032	0.007-1.9	2
ABA	$y = -7E\text{-}07x^2 + 0.0034x - 0.0062^{ł}$	0.9987	0.05994	0.21234	0.01-1.5	3
GA_1	$y = -0.0001x^2 + 0.0293x – 0.0011^{ł}$	0.9995	0.00369	0.01476	0.008-0.5	3
IAA	$y = 6E\text{-}05x^2 + 0.0042x + 0.0002^{ł}$	0.9993	0.00119	0.00351	0.0002-0.3	3
SA	$y = 0.0073x^2 + 0.1006x + 0.0436^{ł}$	0.9994	0.32987	1.33083	0.005-10	3
GA_4	$y = -0.1727x^2 + 0.9958x - 0.0112^{ł}$	0.9982	0.03782	0.1488	0.007-0.9	3
[a]limit of detection - [b]limit of quantification - [ł]blank - [§]matrix						

Table 5.12: Calibration curves for the samples run in blank solvent and receptacle matrix matched solvent.

Calibration curves equations for the prediction of phytohormones amount in the crumbly induction experiments. Data from samples run both in 50% methanol (blank) and in matrix matched (e.g. 50 mg of raspberry receptacle processed and then resuspended in 50% methanol) and obviously related to the sole phytohormones detectable/quantifiable with the analytical methods optimised for these target compounds in this work. All the curves revealed a broad linear concentration range, spanning over three orders of magnitude and had a correlation coefficient (R^2) of at least 0.989. In addition, for each calibration curve was reported the limit of detection (LoD) and the limit of quantification (LoQ).

The calibration model equations were all quadratic since they fitted the data more precisely than the linear functions. This was as a result of the large difference in the target analyte concentration between the two different stages (i.e. green and red fruit). For such reason, for ABA in drupelet (third batch), two different calibration curves were produced, one at the lowest linear range available and the other at the highest (see Table 5.11). This facilitating broad dynamic range to the detection/quantification. The calibration curves were run in three different ways, in pure solvents (blank) and in matrix, respectively of drupelet and receptacle (see section 5.2.7 for further details) and while in general the most performing equations were obtain from the serial dilution run in blank, for some target analytes (i.e. ABA, SA and GA_4), and only in some batches, equations were calculated from the calibration samples run in matrix (see Tables 5.11 and 5.12 for full details).

Twelve samples, in both pure solvent (blank) and matrix matched (i.e. drupelet and receptacle), with specific set concentrations, a consequence of eleven serial dilution from a set starting concentration, were analysed and the instrument response ratio (peak area of target compound divided peak area of the corresponding ISTD) was plot against the specific concentration of the analyte at which that peak area ratio was measured. In this manner a curve with, in theory, up to twelve points was produced for each target analyte. In all cases this generated a polynomial quadratic relationship. Figure 5.11 highlights examples of calibration curves for three out of the five detected/quantified target phytohormones (i.e. GA_1, IAA and GA_4) while Figure 5.12 shows example of calibration curve for the two remaining target compounds (i.e. ABA and SA).

For each of the 120 samples, application of the calibration curves (see Table 5.11 and 5.12), facilitated analyte concentration calculation. Obviously these were selected only for those above the limit of quantification (LoQ), (Tables 5.11 and 5.12). Unfortunately quantification could only be partially achieved as only in some cases was the amount of the analyte in the samples above the LoQ. The only exception was GA_1 whose amount, at least in the receptacle samples, was always above the LoQ (Table 5.13). In Figure 5.13 was reported the bar chart of the relative amount (i.e. ng/50 mg) of sample for the only two phytohormones (i.e. ABA and GA_1) showing differences statistically significant between the two phenotypes (i.e. artificially induced ‘crumbly’ and control); with ABA being significantly higher, in receptacle of artificially induced ‘crumbly’ at green stage, while GA_1 significant lower in drupelet at red stage always of the artificially induced ‘crumbly’ samples. In Figures 5.14 were reported the bar chart of the relative amount (i.e. ng/50 mg) for all the five target phytohormones of the control samples at green stage for

both drupelet and receptacle. The same was reported in Figure 5.15 but this time for the 'artificially incuded 'crumbly' samples. In Figure 5.16, the relative amount (i.e. ng/50 mg) of both drupelet and receptacle at red stage for the control samples; the same was reported in Figure 5.17 but this time for the artificially induced 'crumbly samples. At the green fruit stage, salicylic acid was by far the most abundant target phytohormones in all the samples (Figures 5.14 and 5.15). Salicylic acid together with ABA, GA_1 and GA_4, were measured at a higher concentration in the receptacle in all the samples (see Figures 5.14 to 5.17). IAA was the only target compound whose concentration were always higher in the drupelet than the in receptacle as well as at the green stage rather than red fruit stage; all the other target compounds were always higher in receptacle. At the red fruit stage, in the red drupelet ABA was the most abundant phytohormone (Figures 5.16 and 5.17) while SA was the compound with the highest concentration in the receptacle (Figure 5.16 and 5.17). The two gibberellins (i.e. GA_1 and GA_4) were the only compounds with no statistically significat differences between the two stages (Table 5.10). Again there was no statistically significant difference between the two phenotypes (i.e. artificially induced 'crumbly' and control) except for ABA in the receptacle at the green stage and for GA_1 in the drupelet at the red stage (see Figure 5.13) with ABA concentration being statistically significantly higher in the artificially induced 'crumbly' while that of GA_1 being lower.

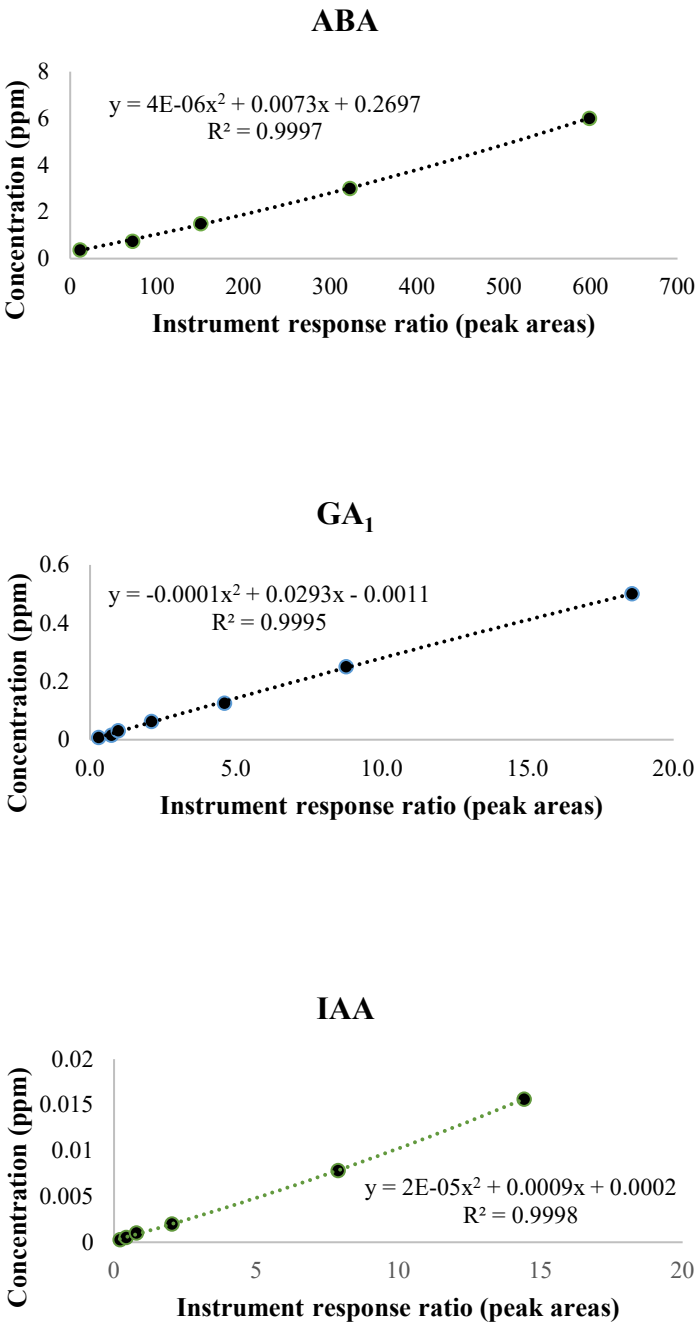


Figure 5.11: Calibration curves in matrix matched or blank solvent for the five detected/quantified target phytohormones.

Examples of three calibration curves related to three of the five target phytohormones (i.e. ABA, GA_1 and IAA) detected/quantified with the LC-MS analytical method described in this chapter. The first one referred to samples run in matrix matched and more precisely to ABA (receptacle) while the other two, GA_1 and IAA referred to samples run in pure solvent (blank). For all these three detected/quantified target phytohormones, the calibration model equations were quadratic because they fitted the data more precisely than the simple linear functions.

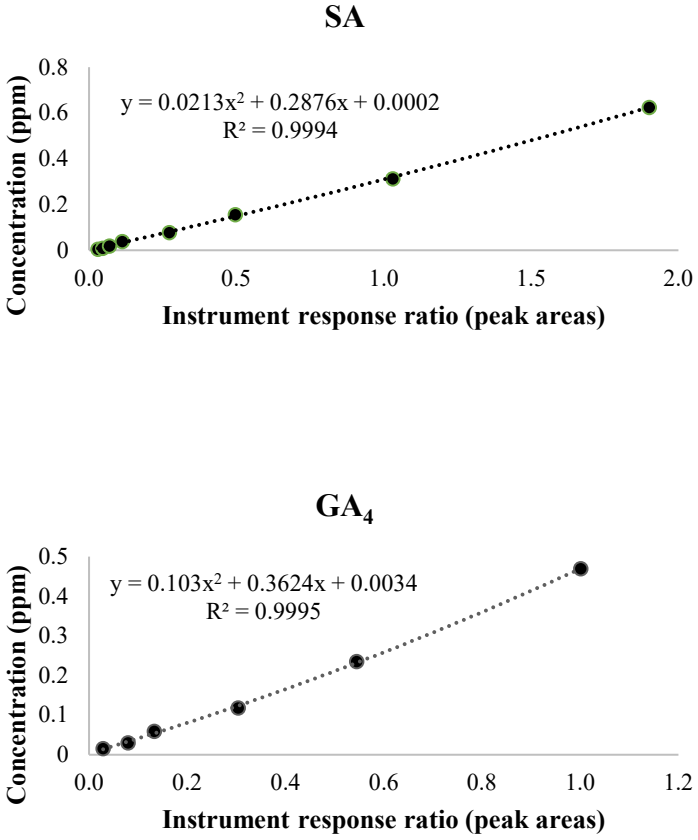


Figure 5.12: Calibration curves in matrix matched or blank solvent for the five detected/quantified target phytohormones.

Examples of other two calibration curves related to the remaining two target phytohormones (i.e. SA and GA_4) detected/quantified with the LC-MS analytical method described in this chapter. Both SA and GA_4 referred to samples run in drupelet matrix matched. As for all the other detected/quantified target phytohormones, the calibration model equations were all quadratic because they fitted the data more precisely than the simple linear functions.

A	compound	type	part (ng)			
			D_green	R_green	D_red	R_red
	ABA	control	0.31±0.07	0.41±0.05[a]	4.69±0.66	7.17±0.81
		crumbly	0.37±0.10	0.94±0.11[a]	4.37±0.49	7.75±1.15
	GA_1	control	0.13±0.03	0.58±0.09	0.20±0.05[a]	0.65±0.09
		crumbly	0.18±0.05	0.77±0.19	0.08±0.01[a]	0.68±0.13
	IAA	control	0.09±0.02	0.03±0.006	0.04±0.01	n.d.
		crumbly	0.07±0.01	0.02±0.003	0.05±0.01	n.d.
	SA	control	0.73±0.09	149.48±33.32	0.46±0.08	38.35±7.52
		crumbly	0.71±0.13	104.83±19.87	0.39±0.07	25.02±5.15
	GA_4	control	0.26±0.05	1.91±0.52	0.24±0.03	1.94±0.58
		crumbly	0.25±0.04	1.33±0.37	0.24±0.03	1.09±0.27
B	**compound**	**type**	**part (ng)**			
			D_green	**R_green**	**D_red**	**R_red**
	ABA	control	2.46±0.60	3.30±0.43[a]	37.6±5.26	57.38±6.45
		crumbly	3.00±0.80	7.56±0.88[a]	35.0±3.90	61.98±9.24
	GA_1	control	1.03±0.25	4.65±0.79	1.62±0.41[a]	5.24±0.72
		crumbly	1.45±0.38	6.18±1.54	0.65±0.10[a]	5.47±1.05
	IAA	control	0.74±0.17	0.24±0.05	0.29±0.05	n.d.
		crumbly	0.53±0.10	0.15±0.02	0.43±0.09	n.d.
	SA	control	5.82±0.71	1195.82±266.57	3.69±0.65	360.80±60.13
		crumbly	5.74±1.04	838.67±158.93	3.12±0.55	200.15±41.17
	GA_4	control	2.08±0.37	15.25±4.17	1.95±0.29	15.56±4.61
		crumbly	1.99±0.31	10.62±2.95	1.93±0.28	8.73±2.15

[a]difference statistically significant at 95% confidence levels
D (drupelet) – R (receptacle)

Table 5.13: Approximative quantification of ABA, GA_1, IAA, SA and GA_4 in receptacle and drupelet at green and red stage and for both control and artificially induced 'crumbly' samples.

Average of the amount and standard error (S.E.) of target compounds in the 5 µL injection volume (A) and its equivalent in 50 mg of sample (B).

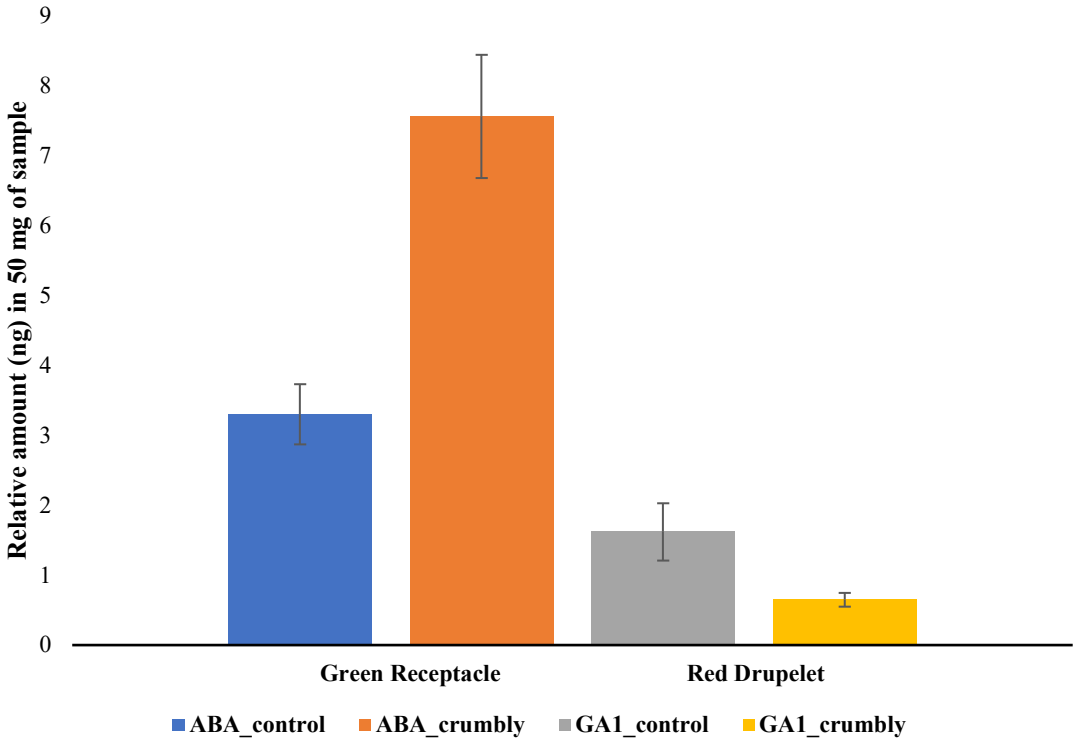


Figure 5.13: Bar chart of the relative amount (ng) of ABA and GA_1.

Bar chart of averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of ABA and GA_1 in control and artificially induced 'crumbly' green receptacle (ABA) and control and artificially induced 'crumbly' red drupelet (GA_1); differences being significative at about 95% confidence levels.

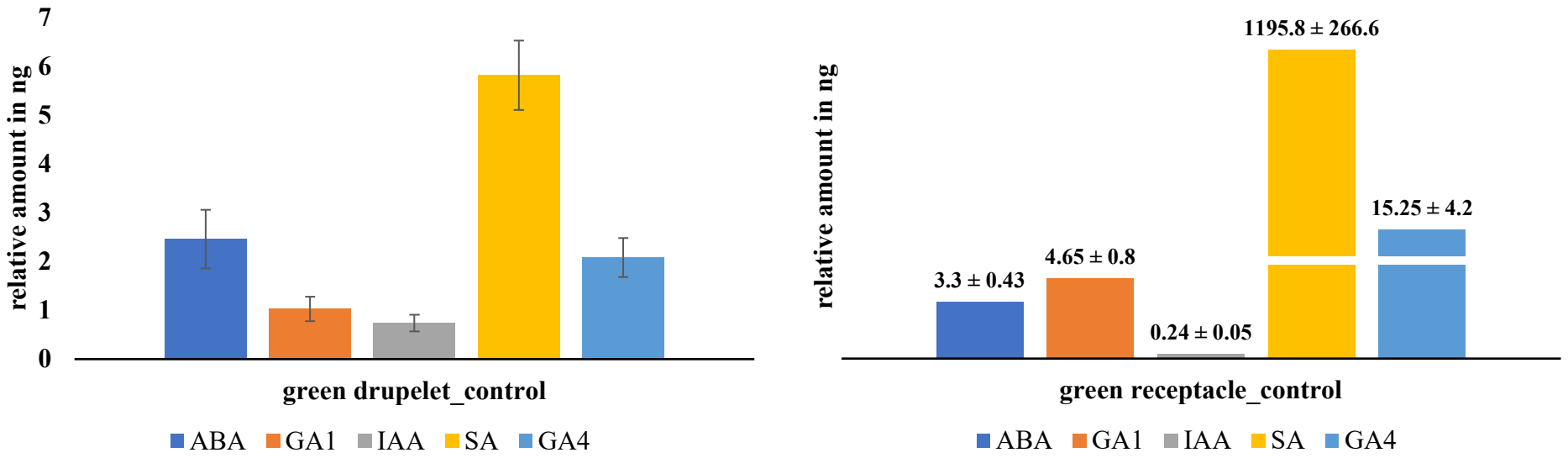


Figure 5.14: Bar chart of the relative amount (ng) of all the five detected/quantified phytohormones in both drupelet (left chart) and receptacle (right chart) at green stage in control samples.

Bar chart of averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of the five phytohormones in control green stage of drupelet and receptacle. Differences between parts being significant at about 95% of confidence levels.

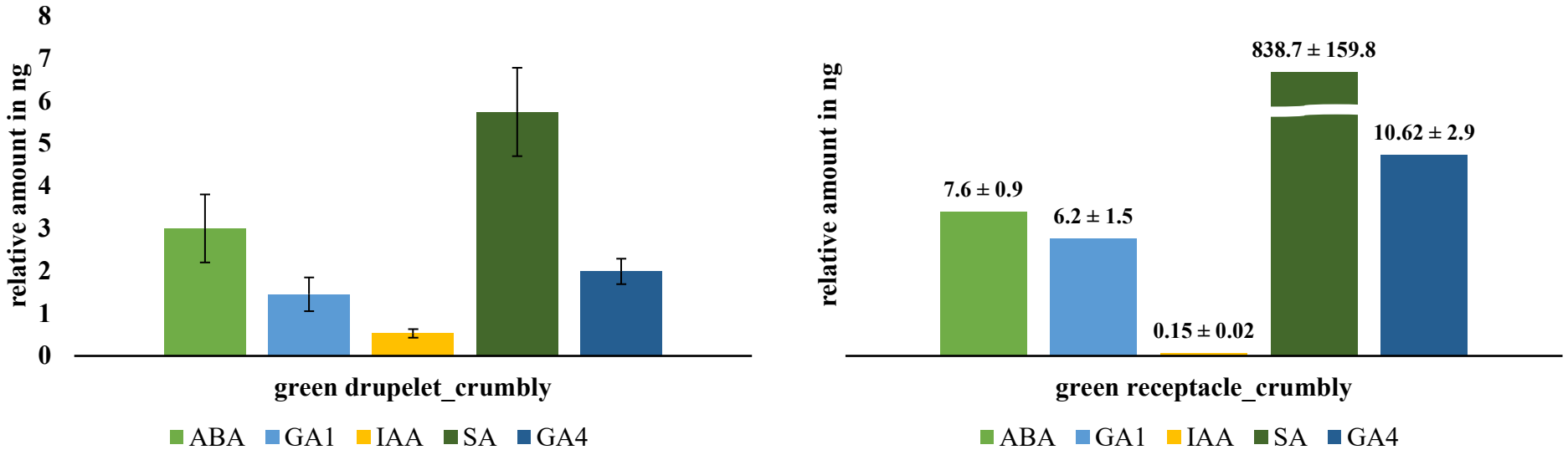


Figure 5.15: Bar chart of the relative amount (ng) of all the five detected/quantified phytohormones in both drupelet (left chart) and receptacle (right chart) at green stage in artificially induced 'crumbly' samples.

Bar chart of averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of the five phytohormones in crumbly green stage of drupelet and receptacle. Differences between parts being significative at about 95% of confidence levels.

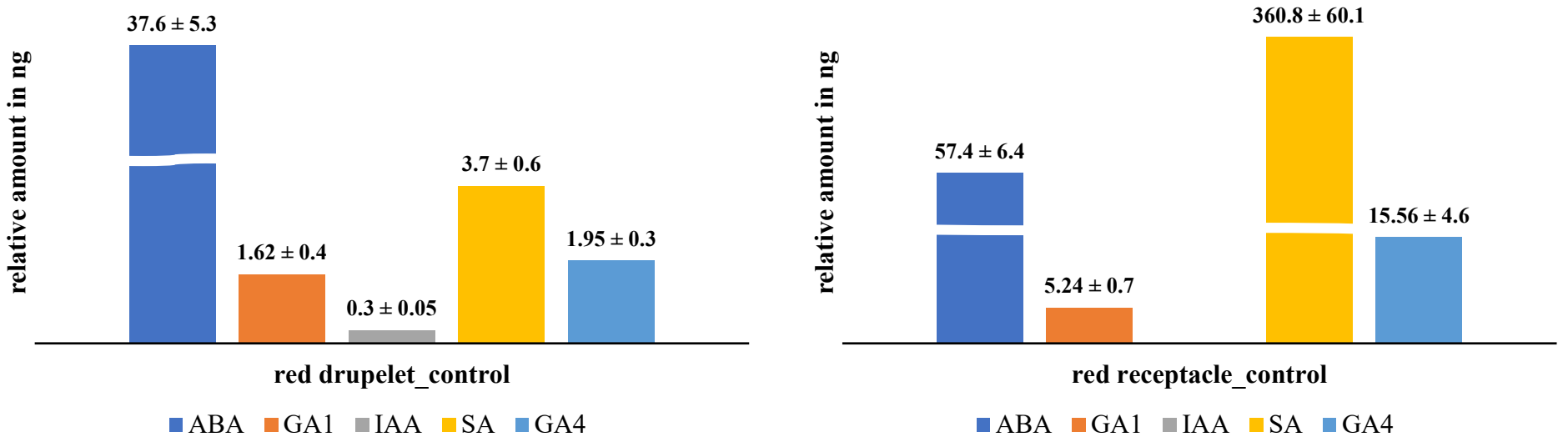


Figure 5.16: Bar chart of the relative amount (ng) of all the five detected/quantified phytohormones in both drupelet (left chart) and receptacle (right chart) at red stage in control samples.

Bar chart of averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of the five phytohormones in control red stage of drupelet and receptacle. Differences between parts being significative at about 95% of confidence levels.

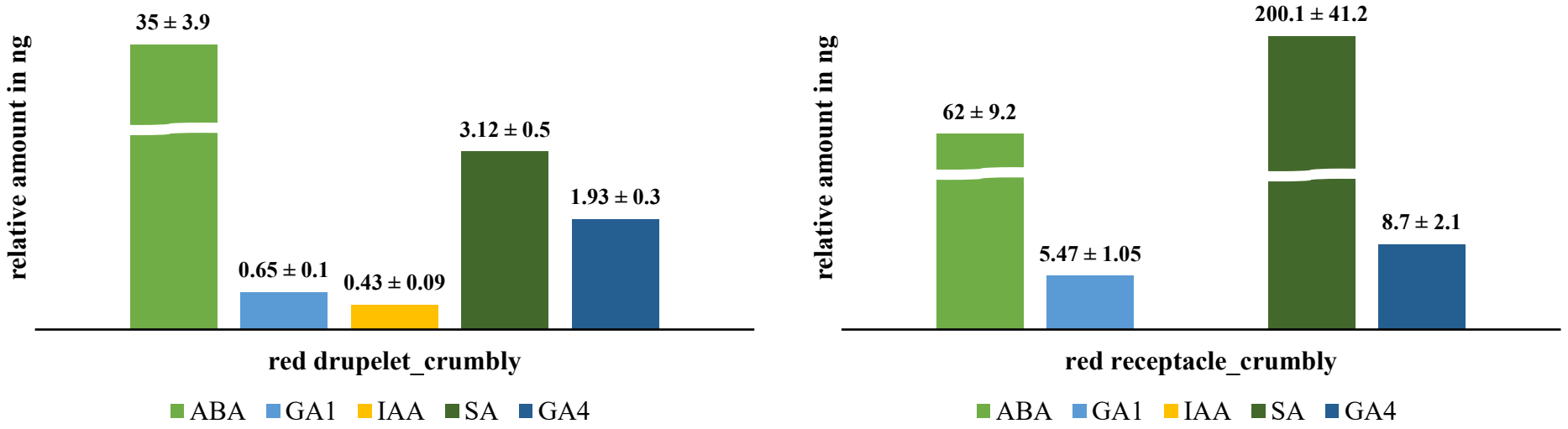


Figure 5.17: Bar chart of the relative amount (ng) of all the five detected/quantified phytohormones in both drupelet (left chart) and receptacle (right chart) at red stage in artificially induced 'crumbly' samples.

Bar chart of averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of the five phytohormones in crumbly red stage of drupelet and receptacle. Differences between parts being significative at about 95% of confidence levels.

In conclusion the metabolomic analysis allowed the partial quantification of the target phytohormones with GA_1 being the only exception where full quantification was achieved and only for the receptacle samples. The most abundant compound by far was SA and its amounts in receptacle were only indicative. It could not be estimated properly since the values were much greater than the linear range of the calibration curves; the dilution and subsequent remeasurement of diluted samples could not be performed due to the extremely unstable nature of these target compounds, rapidly degradable in aqueous solutions. The second most abundant target phytohormone was ABA whose amounts were always greater in the receptacle. The only compound found in higher concentrations in drupelet compared to receptacle was IAA (see Figures 5.13 to 5.15).

5.4 Discussion

It is well established that plant hormones are signalling molecules that participate in the regulation of fruit development processes. Fruit growth is a complex process where the key stage is the pollination/fertilization which triggers cell division and then cell enlargement for both ovary (pericarp) and ovule (seed) development (Ozga and Reinecke, 2003). Phytohormones action and their interactions affect fruit development in different ways. For instance, the biosynthesis of a hormone might be affected by another hormone or the hormonal signal transduction pathway might be affected by an analogous component(s) of a different hormone signal transduction pathway. In other cases, different classes of hormones, affecting the growth of different cells or tissues, must act in concert to guarantee the normal fruit development or a class of hormones begins a process but to complete it, another class of hormones is required (Ozga and Reinecke, 2003).

Many plant hormones groups (auxins, cytokinins, gibberellins, salicylates, jasmonates and ABA) are involved in the fruit development process but a clear model of the hormonal regulation system is not currently available. Fruit growth can be divided into two main phases, before and after pollination/fertilization. During the first phase, the production of new cells is the leading factor responsible for the development of flower and fruit primordia. Low levels of auxins and their conjugates, as well as gibberellins, maintain temporally inactive flower and fruit primordia (Obroucheva, 2014). Growth is slowed and senescence occurs unless events triggered by pollination/fertilization, such as cell division and cell enlargement both in the ovary (fruit pericarp) and in the ovules (seeds), occur to promote the fruit growth (Ozga and Reinecke, 2003). For a review of current knowledge on the stages in raspberry development to fruiting see Graham and Simpson (2018).

The results of the LC-MS analysis on the samples from the crumbly induction experiments, showed that, from the perspective of ‘crumbly fruit’, of the eighteen target phytohormones only two, ABA and GA_1, were significantly different at about 95% confidence level for what concerned the phenotype*stage*part interaction. ABA was significantly higher in green receptacle of artificial ‘crumbly’ samples while GA_1 was significantly higher in red drupelet of normal fruit (control).

Of the five target phytohormones detected/quantified during this work, four of them (i.e. ABA, GA_1, SA and GA_4) were higher in content in receptacle and only IAA was more

abundant in drupelet samples. Moreover, while the amount of the two gibberellins (GA_1 and GA_4) was not significantly different between the two stages, the other three plant hormones were respectively higher at green stage (i.e. IAA and SA) and higher at red stage fruit in the case of ABA.

The growth and development of a complex fruit such as the raspberry one, depend on many factors, anatomy and physiology of the flower, together with the physiological/molecular processes triggered by the pollination/fertilization and indirectly from the environment. The hypothesis behind the design of the 'crumbly' fruit induction experiments was that, to allow the synchronized growth of all ovaries, once fertilised by pollen each ovary would send a signal to the receptacle and once a threshold was reached to allow the formation of a normal size fruit, the receptacle would trigger the simultaneous growth of all the ovaries.

In the artificially induced 'crumbly' fruit, the damage to the receptacle, in theory, should compromise this hypothetic regulating process. In fact, although all the available carpels were manually pollinated and it would be reasonable to suppose that the fertilised ovules would have sent the molecular message, the receptacle being damaged would not be able to exert its normal function. No back signal, or at least only partially, would be sent to the ovules for promoting and inducing their coordinated growth. Such a lack of regulation would cause the independent growth of each of the ovaries with those fertilised earlier growing more than the others, leading to the formation of misshapen fruit with drupelets of different size.

According to the current models, fruit development is regulated by the interplay of different hormones. The combined action of auxins, gibberellins, cytokinins seem to play the main role in fruit set and in fact their levels increase substantially. This has been already validated by exogenous applications of these three hormones which cause parthenocarpic fruit in tomato (Kumar et al., 2014). ABA levels decrease at fruit set as a consequence of downregulation of the genes involved in ABA biosynthesis (NCED1 and NSY) and of upregulation of a gene (CYP707A) involved in its degradation (Vriezen et al., 2008).

The development of fruit and seed is intimately connected and synchronised with the last being a source of hormones, especially auxins, gibberellins and cytokinins, which are then transported in the surrounding tissues where they are involved in stimulating the growth and even determine the final size of the fruit (Ozga and Reinecke, 2003).

A confirmation of this latent model might be found in tomatoes where, the inactive state of the ovules is maintained by high concentration of ethylene and ABA (Vriezen et al., 2008, Obroucheva, 2014). After pollination/fertilization, in the fertilized ovules of tomato, two main events occur, the reduction of ethylene and ABA and an increase in the content of auxins. In particular, the main event leading to fruit set is the increase of auxin levels because they induce the biosynthesis of gibberellins. Both these two classes of phytohormones promote active growth of fruit but while auxins influence cell division, the gibberellins are responsible for cell extension (Serrani et al., 2008). The auxin IAA was detected/quantified with the LC-MS method and the results of the analysis showed that its concentration, at early stage of fruit development (green stage), was higher in drupelet compared to receptacle, suggesting its role as potential candidate to assist growth of fertilised ovaries. The principal component analyses supported this latent function in the drupelet since in the first principal component, the IAA vector was the only one with a negative direction, consistent with its higher content in drupelets compared to the other four positive vectors, whose amount, in antithesis, were higher in receptacle. Further considerations might be proposed by examining the interactions between hormones. Auxins tend to act downstream of ABA when the two hormones regulate the same process and this, regardless of their effect (i.e. same or opposite) on that process (Emenecker *et al.,* 2020). Studies seem to suggest that ABA may promote auxin activity in one context, when both hormones have the same effect, or repress auxin activity in cases each hormone has opposite effect on the process (Emenecker *et al.,* 2020). Gibberellins (GAs) whose biosynthesis is induced by auxins (Serrani *et al.,* 2008), like the latter, might interact antagonistically with ABA as occur during flowering with GAs having a positive effect on flower transition while ABA having both negative and positive effect (Shu et al., 2018); although further investigation is required, it cannot be excluded, at early stage of fruit development a similar antagonistic GAs-ABA interaction.

In *Capsicum annuum* IAA appears to be responsible for the formation of vascular connection between the ovule and the rest of the carpel that contributes to the sustained growth of the fruit (Tiwari et al., 2013). This would explain the higher level of IAA at green stage in the drupelets. Auxins in general, IAA in particular, induce gibberellin biosynthesis that contributes to cell expansion (Ozga et al., 2003). Gibberellins were detected/quantified in both drupelet and receptacle but with higher levels in receptacle samples and with no significant differences between the two stages (i.e. green and red fruit). The PCA analysis at green stage showed that GA_1, followed by SA, was the main

vector, the one with higher magnitude and then the one that contributed the most to the sample separation highlighting then, the pre-eminent role of GA_1 in receptacle during early stage of fruit development. This supports the role of IAA promoting cell division and the formation of vascular connections in the drupelets and GA_1 stimulating cell expansion in the receptacle. Salicylic acid (SA) was the second highest vector in magnitude, at green stage. Salicylic acid is usually associated with response to abiotic and biotic stresses and with systemic acquired resistance (SAR) but it is an important component in signal transduction systems (Elwan and El-Hamahmy, 2009). IAA and SA share a common precursor (chorismate) and crosstalk between them has been shown to occur during fruit development and ripening (Pérez-Llorca *et al.,* 2019). Studies conducted on tomatoes showed that IAA interact with SA through the auxin response factor (ARF2) whose overexpression cause the reduction of SA levels suggesting a potential role of IAA in limiting SA during fruit development (Pérez-Llorca *et al.,* 2019). Studies on the kernel of *Zea mays* suggests an important modulating effect of SA on IAA (LeClere et al., 2008) and so in accordance with the results of the targeted phytohormones analyses, a model for fruit development in raspberry, at early stage, would envisage crosstalk between four hormones, GA_1, SA, IAA and ABA. Afterwards ovule fertilization, an increase in IAA level in the ovary would cause reduction of ABA, promote cell division and increase GA_1 levels. SA, presented in higher amount in the receptacle would modulate IAA, keeping its level low, and in fact IAA was significantly lower in receptacle, and this effect should probably limit cell division but favour cell expansion since the levels of GA_1, as reflected in the targeted phytohormones analysis, were significative higher in the receptacle than in drupelet. In addition, and always in the receptacle, SA inhibiting IAA would allow to keep ABA levels higher than in drupelets.

In raspberry, this proposed early stage fruit growth model with four main drivers (i.e. GA_1, SA, IAA and ABA) did not seem to explain fully the situation found with the analysis of the 'crumbly' induction samples where, at the green stage in the receptacle of the artificial 'crumbly' samples, a significantly higher level of ABA, compared to the receptacle from control samples, was detected/quantified. A possible explanation for this could be in the early stage of fruit development, of a further class of phytohormones, the cytokinins (CKs) have a role. CKs function in receptacle would be to replace IAA and attenuate ABA level, since they act differently to IAA. CKs would not be modulated by SA and then could exert their role of limiting ABA. Cytokinins also regulate cell division in both pericarp and seed. For instance, during the first 10 days of fruit growth in lupine,

cytokinins levels increase in the pedicel due to intense import through phloem and xylem vessels. Such accumulation might be related to downregulation, induced by GAs, of genes encoding enzymes involved in cytokinins conjugation such as zeatin glucosyl transferase and zeatin xylosyl transferase (Vriezen et al., 2008). Studies in *Capsicum annuum* (Honda et al., 2017, Tiwari et al., 2013) showed the important role exerted by cytokinins, in concert with gibberellins, in early stage of fruit development, with respect to predominantly cell enlargement. In normal fruit, high level of cytokinins in the receptacle at green stage would modulate ABA level and promoting cell expansion. Another important aspect to be considered is the stage of fruit development at which the analyses were performed. The phase (early green fruit) to which the samples were collected and analysed corresponds approximately to two weeks after anthesis and at that stage, the cell divisions processes could be already ended, for instance in cucumber its blast fulfils within four days after anthesis (Boonkorkaew et al., 2008); much earlier than the two weeks of the samples analysed in this work. For such reason, answers concerning the initial stages of fruit growth/development, related to any latent hormonal fruit growth regulated processes cannot be drawn from the experiments simply because these might be very probably already occurred.

For reasons requiring further investigation, in artificial 'crumbly' fruit, at the early stage of fruit development, a lower content of CKs in the receptacle would cause a disequilibrium in the phytohormones content, in particular a higher level of ABA, that would impede the receptacle to carry out its function of regulating and synchronising the growth of all the fertilised ovaries. In the artificial crumbly fruit all the carpels were hand pollinated so it would be expected they would develop fruits with the maximum number of drupelet but damage at the receptacle, prior to pollination would be responsible for hormonal disequilibria such as the one identified above with higher level of ABA in the green receptacle that would cause the development of misshapen 'crumbly' like fruit with reduced number of drupelets.

Fruit ripening involves the activation of several genetic and biochemical pathways leading, in a coordinated and synchronised fashion, to colour and chemical changes, conversion of complex carbohydrates to sugars, accumulation of flavour and aroma compounds and change in the cell wall dynamics that cause either dehiscence or softening. Ethylene seems to be only weakly involved during fruit reddening as the upregulation, in strawberry, of genes coding ethylene receptors. Auxins, while involved during fruit growth and maturation, start to decrease during the white fruit stage causing the suppression of genes encoding endo-β-glucanase, pectate lyase and pectin

methyltransferase as well as of genes involved in the final conversions of flavonoids and in the synthesis of aromatic compounds (Harpster et al., 1998).

At the red stage, the results of the targeted phytohormone analysis showed again that, five compounds were detected/quantified, with four more abundant in receptacle (ABA, GA_1, SA and GA_4) and only IAA more abundant in drupelet. Compared to the green stage, except for ABA which increased and gibberellins that remained stable, both SA and IAA were reduced in the late stage of fruit development. The PCA analyses showed that SA, ABA and GA_4 were the vectors with the highest magnitude and they were negatively correlated with IAA vector which was the only one having negative direction. In fact, IAA was the sole phytohormone with higher level in drupelet. The increase of ABA compared to the green stage would be coherent with its leading role in fruit ripening. In fact, regarding the hormonal regulation of fruit ripening, in non-climacteric crop such as raspberry, ABA seems to be the only phytohormone directly involved in this process. In strawberry, for instance, treatments with 0.5 mM ABA turned white mature berry into deep red berries within a week while untreated berries did not reach even pink colour in the same time period. The main role of ABA in fruit ripening might be confirmed by upregulation of genes encoding ABA precursors such as xanthoxine as well as ABA receptor such as PYR1. Experiments in tomato, aiming to delay ABA accumulation, caused a delay in fruit ripening (Zhang et al., 2009). In peach exogenous application of ABA increased the carrier-mediated transport through both tonoplast and plasma membranes suggesting an important role of this hormones with respect to sugar accumulation in fruits (Kondo et al., 2004).

Target gibberellins levels remained constant between the two stages and the analysis of the variance (ANOVA) for the crumbly induction experiments showed that GA_1 was the sole target plant hormones whose differences, between artificial crumbly and control samples, were significant at about 95% of confidence levels in the red drupelet. More specifically the level of GA_1 in control samples was significantly higher than that in artificial crumbly red drupelets samples. No other target hormones were found to be significantly different between the two phenotypes (i.e. artificial crumbly and control) and so to formulate a fuller explanation of the phenomenon further hormones need to be considered. Cytokinins would be interesting candidates, at maturation, when fruit prepares to undergo ripening, auxins and cytokinins appear to be involved and in fact in tomato ripening inhibitor (rin) mutants, the levels of both these two hormones were higher at maturation stage than those of wild-type fruits (Davey and Vanstaden, 1978). In

strawberry the removal of achenes, the source of auxins from the surrounding tissues, causes rapid ripening of the receptacle unless exogenous applications of auxins stall the process (Given et al., 1988). Auxins seem to play the major role at fruit maturation, observations in strawberry and mango suggest that reduction of auxins represents a prerequisite to commence ripening (Zaharah et al., 2012, Given et al., 1988). However experiments with a cytokinin-deficient mutant in *Arabidopsis thaliana* showed unsynchronized ripening phenotype (Kumar et al., 2014) while observations of cytokinins decreasing before ripening initiation suggest that cytokinins too might play a role in fruit maturation (Bottcher et al., 2011).

On the basis of the targeted phytohormones analyses results, SA, the vector with highest magnitude, modulating IAA would allow ABA levels to increase while gibberellins, being not affected by SA, would remain constant. In drupelets, the crosstalk between SA and IAA appears to lead fruit maturation in raspberry normal fruit; this would find evidence from experiments on papaya. Exogenous application of SA caused altered expression of IAA genes with some being down-regulated while others up-regulated (Liu et al., 2017a) and although further research is required, an indication of the presence of a cross-talk between IAA and SA during fruit ripening seems clear.

The phytohormones analysis at red stage fruit showed that in the drupelet of artificial 'crumbly' samples the level of GA_1 was significantly lower compared to normal fruit (control). In such a scenario the action of other classes of hormones besides those analysed from here will be required to give a fuller explanation of the process. Cytokinins could be good potential candidates here. Supporting this, cytokinins too participate in varying steps of fruit growth as previously studied by, for instance, Kumar et al., (2014) or Davey and Vanstaden (1978) and so it is plausible to theorize about their involvement in fruit maturation too. Unlike IAA, cytokinins do not experience modulation from both SA and ABA, and so they could cause hormonal disequilibria that might involve other class of hormones besides those found here that for reasons needing further investigation might cause in the drupelet of crumbly phenotype samples at red stage an increase in GA_1 level.

The gene expression analysis of the Glen Moy x Latham progeny (see chapter 3 and 4 for further details) showed the upregulation in the mostly 'crumbly' phenotype, at green stage of three genes involved in cytokinins biosynthesis/metabolism and in response to. All these three genes, differently expressed between the two phenotypes (i.e. mostly and never 'crumbly'), were mapped in the two 'crumbly' QTLs on linkage group 3 and more

precisely on gene in the original 'crumbly' QTL and one in the QTL identified here in this work (see chapter 4 for further details). These findings provided further evidence of the possible part played by this class of hormones (cytokinins) in fruit development process.

In summary, two hormonal fruit growth models, one at early fruit development (green fruit) and the other at late stage (red fruit) were suggested. On the basis of the results of the targeted phytohormones analyses in crumbly induction experiments (see section 5.2.1) at the early stage of fruit development, the crosstalk between the two fruit tissues analysed (i.e. receptacle and drupelet) involving five hormones gave a general idea of fruit development regulation. In this IAA level increased in drupelets after fertilization, leading to ABA modulation and then break of the inactive state controlled by high levels of ABA. IAA stimulated gibberellins production but, while in receptacle, the high level of SA modulated IAA keeping its level low in this flower part. The same effect would be exerted on GA_1 whose levels were higher in receptacle. Such a scenario would see the promotion, primarily, of cell division in drupelets mediated by IAA and cell enlargement, mediated by GA_1, in receptacle. The significantly different level of ABA in the green receptacle compared to control could be explained by proposing the interplay of a further hormone, not detected in the analyses here, the cytokinins that would modulate ABA in the receptacle since their effect, as opposed to IAA, would not be affected by SA. It might be postulated that in the green receptacle of normal fruit the relative hypothetical higher level of cytokinins would keep ABA relatively lower creating the optimum conditions for normal fruit growth, for instance by promoting cell expansion, acting in concert with GA_1. In artificial crumbly samples at the green stage, in receptacle, ABA higher levels could be explained by lower levels of cytokinins that, failing to modulate ABA, would create hormonal disequilibrium responsible for impaired growth. For instance, a reduced degree of cell expansion in the receptacle, leading to misshapen 'crumbly' like fruits.

At the late stage of fruit development, as would be expected for non-climacteric fruits and as the results of the phytohormones analyses showed, ABA levels increased, in both parts (i.e. drupelet and receptacle) compared to green stage while IAA and SA, in drupelets and receptacle respectively decreased. Only gibberellins remained constant in respect to the earlier stage but once again they were significantly higher in the receptacle. In normal fruit the crosstalk between these three hormones (ABA, SA and IAA) would lead the fruit growth. The PCA principal component analyses (PCAs) suggested that the main role was played by SA which modulated IAA, allowing an increase in ABA and creating the ideal

conditions for fruit ripening. The significantly lower levels of GA_1 in red drupelet of artificially induced 'crumbly' samples might be explained by postulating again the involvement of cytokinins, a class of hormones not detected here. The hypothesis would be that the interplay of cytokinins would cause hormonal disequilibrium such as the reduction of GA_1 in drupelet, with the effect of delayed cell expansion that might be confirmed by further analyses. Observations of artificially induced 'crumbly' like fruit at red stage showed a peculiar fruit growth and ripening. In fact, berries took longer to reach their final size compared to normal fruits and moreover while for the latter the transition between red fruit and over ripening lasted few days, in the artificially induced 'crumbly' like fruits the same process took less than one day with fruits reaching full red stage in the afternoon were found overripened early in the morning the following day. Therefore, a much slower fruit enlargement followed by a much faster fruit ripening were the clearest conclusion to be deduced.

5.5 Conclusions

The targeted phytohormones analyses allowed the partial quantification of five compounds: abscisic acid (ABA), gibberellin A_1 (GA_1), indole acetic acid (IAA), salicylic acid (SA) and gibberellin A_4 (GA_4). At green stage, in drupelet and receptacle, the most abundant compound was by far SA, accounting for respectively 42 and 98% of the total amount of the targeted phytohormones studied in this work. ABA, and the two gibberellins were more abundant in receptacle while only IAA was more copious in drupelet (see Table 5.14 for complete semi-quantification data). The diagrams in Figures 5.17 and 5.18 show the relative percentage amounts of the five target phytohormones respectively in drupelet and receptacle at green berry stage.

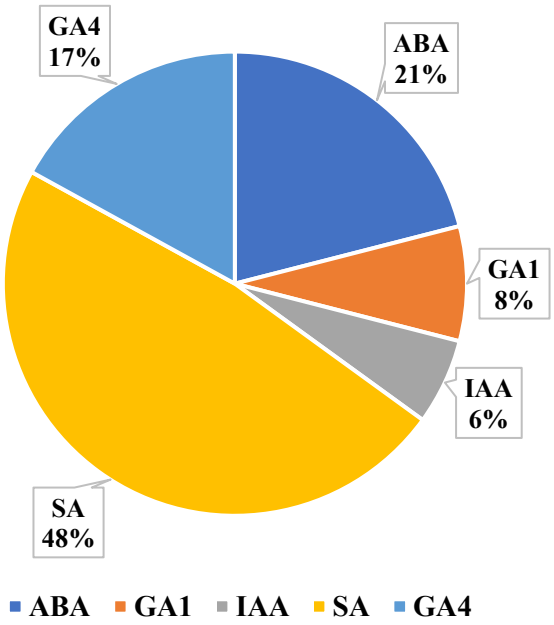


Figure 5.18: Pie chart of the relative amount for the five target phytohormones in drupelet samples at green stage.

Percentage of the relative amount of the five target phytohormones semi-quantified in this work; the amounts refer to predicted values in 50 mg of sample (drupelets at green berry stage).

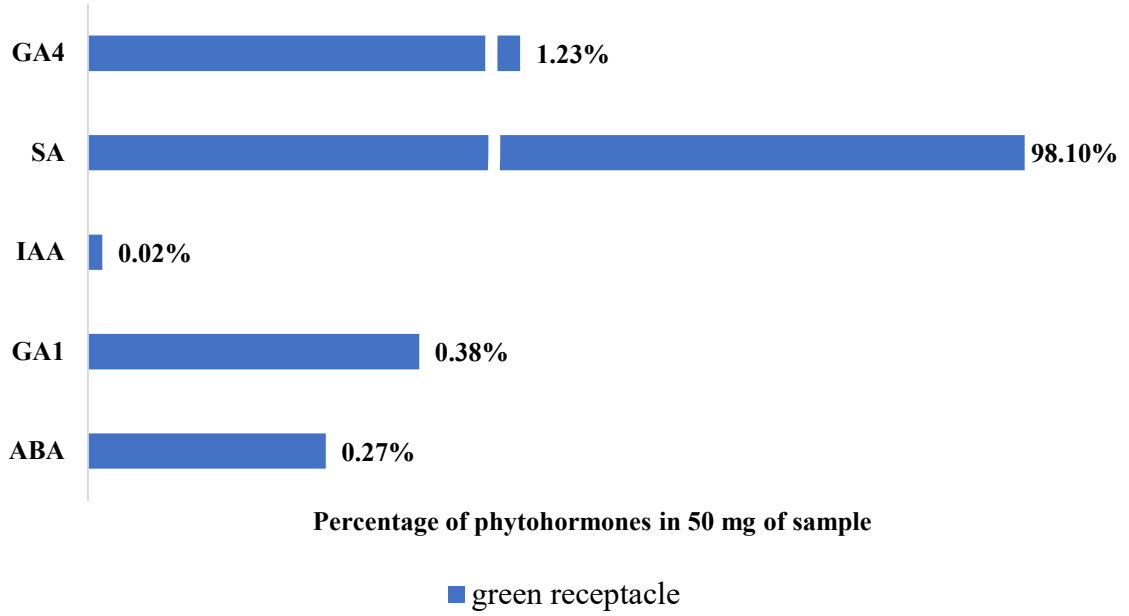


Figure 5.19: Bar chart of the relative amount for the five target phytohormones in receptacle samples at green stage.

Percentage of the relative amount of the five target phytohormones semi-quantified in this work; the amounts refer to predicted values in 50 mg of sample (receptacle at green berry stage).

At red berry stage, SA was still the most abundant compound in receptacle followed by ABA which was, and by far, the most copious phytohormones in drupelet. Once again, the gibberellins were more abundant in receptacle while IAA undetected in this fruit part, was present only in drupelets (Table 5.14). In Figures 5.19 and 5.20 were reported the relative percentage amounts in 50 mg of sample of the five targeted phytohormones at red berry stage.

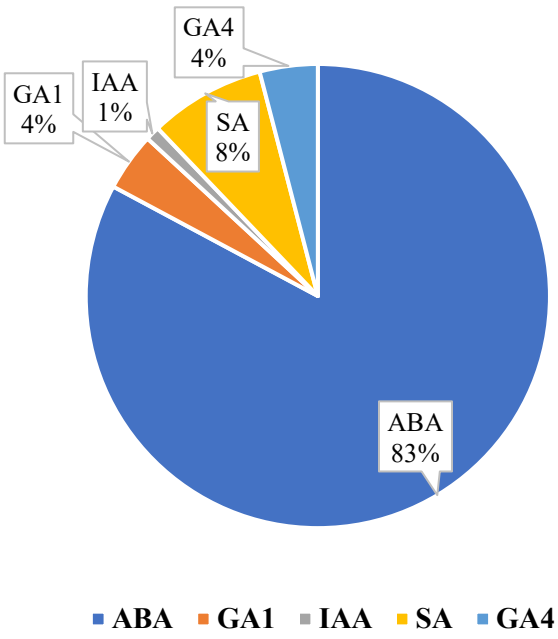


Figure 5.20: Pie chart of the relative amount for the five target phytohormones in drupelet samples at red stage.

Percentage of the relative amount of the five target phytohormones semi-quantified in this work; the amounts refer to predicted values in 50 mg of sample (drupelets at red berry stage).

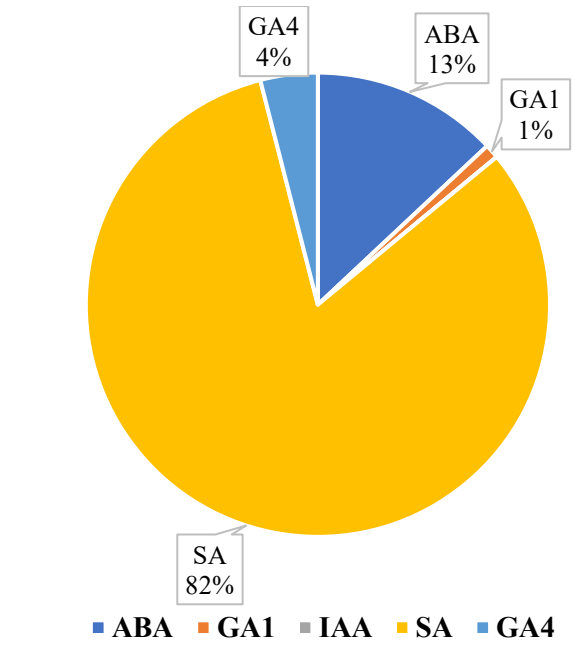


Figure 5.21: Pie chart of the relative amount for the five target phytohormones in receptacle samples at red stage.

Percentage of the relative amount of the five target phytohormones semi-quantified in this work; the amounts refer to predicted values in 50 mg of sample (receptacle at red berry stage).

It is worth specifying that the analysis of variance demonstrated that, except for the gibberellins whose differences were significant only between parts, for all the other three phytohormones these were significant at both the interactions of stage per part. From the perspective of the 'crumbly' fruit induction experiments, the most important finding was the measurement of significant differences between the two phenotypes (i.e. artificially induced 'crumbly' and control) for two targeted phytohormones (Figure 5.23), ABA significantly higher in receptacle at green fruit stage of artificially induced 'crumbly' samples and GA_1 significantly lower in the sample of the same phenotype but in drupelet and at red stage. (see section 5.3.2 for all the details about the results of the ANOVA).

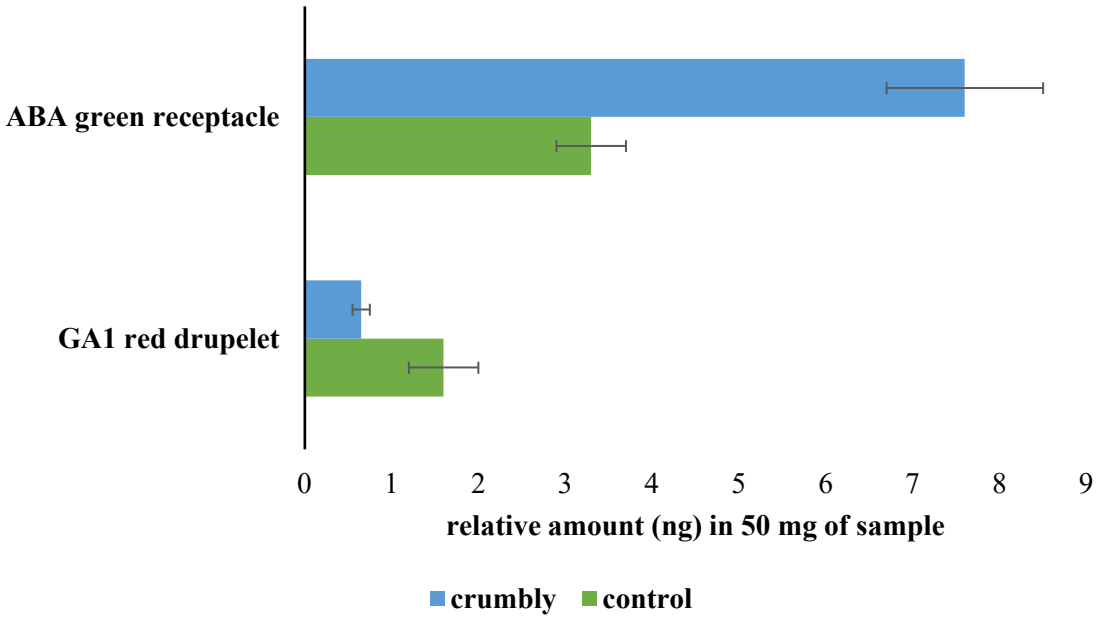


Figure 5.22: Bar chart of the relative amount for ABA and GA_1 between the two different phenotypes (i.e. 'crumbly' and control).

Averaged amount (ng) and standard error in 50 mg samples (average of 5 plants with three biological replicates each plant) of ABA and GA_1 in control and crumbly green receptacle (ABA) and red drupelet (GA_1); differences being significative at about 95% confidence levels.

The principal component analyses (PCA) were performed on samples on both drupelets and receptacle at the two different stages (see sections 5.3.3 and 5.3.4 of this chapter for more details) and they helped identifying the compounds having the major effect on the growth of the two different parts (i.e. drupelet and receptacle) at both the two different stages (i.e. green and red fruit). The results of the analyses allowed to hypothesize two models of fruit parts growth for the two different stages the interplay between these five

hormones measured here plus the introduction of an extra class of growth regulators (cytokinins), that were proposed to better fit the models. This class of plant hormones, contrary to auxins (IAA), are not affected by SA so they can modulate ABA even in presence of high amount of salicylic acid. It can be postulated that cytokinins (CKs), in receptacle at green stage, in relatively high concentration would regulate ABA level by keeping it low to create an optimum status for normal fruit growth; the significantly higher level of ABA, in the artificially induced 'crumbly' samples, not counterbalanced by relatively high values of CKs, in receptacle (green fruit) would explain the hormonal disequilibrium causing the development of misshapen fruits. A similar speculation can be offered for explaining, always in the artificially induced 'crumbly' samples, the significantly lower level of GA_1 in the drupelets but at red stage. Here the conceivable relative high concentration of CKs might be responsible for hormonal disequilibrium driving the reduction of GA_1 levels as measured in those samples.

Chapter 6: General discussion

6.1 Summary

This work set out to develop some understanding of the process leading to crumbly fruit formation in red raspberry. To do this a clear definition of the condition had to be developed. Once the type of crumbly fruit for research was determined a process of study was developed. This involved looking at gene expression and the genetic loci controlling the process and then it developed a process to artificially induce the crumbly phenotype so the hormone profile could be evaluated. Two types of material were used. A field grown mapping population known to segregate for the crumbly condition across seasons and high health stocks of Glen Ample, a cultivar known to express the condition in some seasons and also to be amenable to growing well under controlled conditions. This work identified genes that were differentially expressed between the two phenotypes, confirmed 2 existing QTLs and identified a third new one. A process to induce the 'crumbly' phenotype was developed that allowed the study of the changes in hormone profiles between phenotypes. All of these were central to the proposed hypothesis of a receptacle controlled system regulating and coordinating all those processes occurring soon after fertilization that would represent a key step for the correct growth and development of fruits in red raspberry.

6.2 Crumbly phenotype classification

'Crumbly' fruit is a phenomenon, affecting raspberry plants, that causes the development of misshapen fruits. These uneven fruits when harvested crumble and practically cannot be commercialised, this results in big yield loss for growers. In chapter 1 of this work two new definitions were introduced with the aim of clarifying the discussion about this complex disease/disorder that arises in two different forms.

The Crumbly Fruit Condition (**CFC**), is genetically controlled so can be compared to a genetic illness while the other form, the Malformed Fruit Disorder (**MFD**), is environmentally regulated, and appears more like a disorder (see section 1.3 for more details). Both **CFC** and **MFD** plants show the same symptoms, what distinguish them is the degree to which such symptoms are displayed. Plants affected by **CFC** always produce only misshapen fruit, year after year. In contrast, plants with **MFD,** in general, display symptoms only at the beginning of the fruiting season and only in some seasons, occasionally **MFD** plants produce misshapen fruit, along the whole of a given season but importantly this never happens year after year.

For those raspberry plants designated as having **CFC**, it is apparent that genetic alteration, probably due to somatic mutations in fruit development, cause the formation of aberrant fruits with no commercial value (Jennings, 1988). For those plants affected by **MFD**, the argument is much more complicated; a genetic component is only part in the process, the main driver triggering the condition appears to be the environment. **MFD** causes serious problems for the raspberry industry worldwide.

6.3 Potential triggers

In the introduction of this thesis (see chapter 1), a list of all the potential triggers of 'crumbly' fruit was reported, with nine main factors being listed. Three of these factors (i.e. virus infections, extensive tissue culturing and genetic predisposition) apply for Crumbly Fruit Condition (CFC), the genetic related form of 'crumbly' fruit, while all these nine factors can be linked with Malformed Fruit Disorder (MFD), the more environmental related form of 'crumbly' fruit. These nine factors could be split in four groups.

Those included in the first group were factors impacting early developmental processes(e.g. failed pollination, damage of the flower and excess of nectar) because these situations interfere physically with the fertilization process they could contribute to the formation of 'crumbly' like misshapen fruit. The second group contained plant physiological related factors (i.e. insufficient chilly requirements and adverse environment conditions during key stages of flower/fruit development). Raspberry plants require adequate exposure to chilling temperature (i.e. 0 – 7 °C) indispensable to break bud dormancy (Heide and Sonsteby, 2011) as well as exposure to low temperature and short day-light at the end of the fruiting season to favour the induction of floral initiation and bud dormancy (Woznicki et al., 2016). When these conditions are not met, there are higher risks for the plants to develop anomalous flowers with lower number of carpels which could translate to the production of fruits with lower number of drupelets with consequent higher chances to be misshapen.

The third group comprises pathogen related factors such as viruses and phytoplasmas. In a study conducted on raspberry, cultivar Meeker, Quito-Avela et al., (2014) demonstrated how infections with Raspberry Bush Dwarf Virus (RBDV), Raspberry Leaf Mottle Virus (RLMV) and Raspberry Latent Virus (RpLV), alone or in all possible combinations, were responsible for very severe forms of 'crumbly' fruits. Regardless of the importance of these findings, it is important to clarify that not all plants affected by 'crumbly' fruit are infected with virus and that the same assertion is valid for those with phytoplasmas (see chapter 1 for more details).

The fourth group of potential factors triggering 'crumbly' fruit includes genetic predisposition, with some varieties known to be more prone to the condition than others; an example can be found on Table A.4.1 (see appendix) where a list of sixty-three

different genotypes were reported of which forty-five were labelled as ‘crumbly’ and eighteen as non ‘crumbly’.

The outline of the situation, with respect to those factors that could potentially cause Malformed Fruit Disorder (**MFD**), appears quite complicated with many and different key factors likely to be involved. A detailed study of all potential triggers would not have been possible due to time factors and for that reason this work focused on the study of the underlying mechanism in the plant that lead to the condition and examined both the genetic aspects related to the seasonal triggering of ‘crumbly’ fruit and the potential mechanism in relation to hormone control in the crumbly phenotype that arises due to damage of the receptacle.

6.4 Fruit development

The complexity of the structure of the raspberry fruit suggests that a highly regulating mechanism, both temporal and spacial, is set in place during fruit growth and development. Any anomalies occurring, even in the smallest component/process involved in this complicated apparatus/machinery could explain, in part, the onset of aberrant fruits.

The raspberry fruit is a berry formed by the aggregation of multiple fertilised ovaries (Jennings, 1988). In the flower of raspberry, the number of carpels indicates potentially the number of drupelets forming the final berry; in general, at least 75% of all the carpels must be fertilised to produce a normal fruit. Even in case of small berries, the minimum number of drupelets required to form a regular shape fruit, on average, is sixty. In general, some of the carpels get fertilised via self – or auto-pollination, with this being limited to the outermost carpels whose stigmas receive pollen from the surrounding ring of stamens (see Figure 1.1 in Introduction). For the morphological structure of the flower, the anther on the stamens cannot be in contact with the centremost carpels precluding the complete self-fertilization (Cane, 2005).

On the basis of the above considerations, it would be difficult to anticipate the simultaneous fertilization of all the carpels; on average a minimum of sixty, while it would be more reasonable to assume that plants probably rely on a system that regulates and synchronise the growth of all these many fertilised ovaries. Without this coordination, the earliest fertilised carpels would grow more than those pollinated later, with potential for out of phase development of the drupes and the consequent formation of misshapen fruits.

6.5 Hormonal cross talk hypothesis

In both chapter 1 and chapter 5 of this work, a hypothetical fruit growth regulating process was described with the receptacle acting as central hub. By means of a hormonal crosstalk that would coordinate and synchronise growth and development of all the fertilised ovaries (see chapters 1 and 5 for more details), the formation of normal shape berries could be fulfilled.

An experiment was designed to interfere with this proposed fruit growth regulating process; mechanical damage of the receptacles was performed, soon after emasculation, on flowers of raspberry cultivar Glen Ample grown under optimal conditions. Two days after the treatment, when the stigmas of the carpels became receptive to pollen, the emasculated flowers were hand pollinated (see chapter 5 for details). The hormone content was then measured to identify differences between the artificially induced 'crumbly' and the control samples for both drupelets and receptacles at two different development stages (i.e. green and red berry). A method, described in chapter 5 of this work, was specifically developed for the absolute and simultaneous quantification of eighteen different target plant hormones belonging to six groups of important phytohormones (i.e. ABA, cytokinins, salicylates, jasmonates, auxins and gibberellins).

The analysis proved to be successful in the identification of five of the eighteen target phytohormones, abscisic acid (ABA), gibberellin A_1 (GA_1), indole-acetic acid (IAA), gibberellin A_4 (GA_4) and salicylic acid (SA). Of these target compounds identified, only ABA and GA_1 were found to be significantly different, at 95% confidence levels, between the two different phenotypes (i.e. artificially induced 'crumbly' and the control). ABA was significantly higher in the receptacle at the green stage while GA_1 was significantly lower in the drupelet at the red stage, in the artificially induced 'crumbly' samples compared with the controls.

In chapter 3, the tree cluster heatmap analysis highlighted eight microarray probes, of which five clustered together, matching genes having gene ontology annotation 'response to ABA'. The analysis (see chapter three for details) was performed on the microarray probes, differently expressed between the two phenotypes (i.e. mostly and never 'crumbly) with difference being significant in all the three stages tested (i.e. closed bud, open flower and green berry). These five probes were all upregulated in the mostly 'crumbly' plants and it can be expected that these fruits presented a higher level of ABA

that would explain the upregulation of the five genes activated in response to ABA stimuli. This data presents analogy with the results of the metabolomic analysis where in the receptacle of the green fruit in the artificially induced 'crumbly' samples a significant higher level of ABA was measured. Both these results would indicate an important role played by this hormone in the formation of 'crumbly' like misshapen fruits. In the current models of fruit development, ABA is involved at the beginning of fruit development process where its higher levels help to keep the inactive state of the ovule (Vriezen et al., 2008, Obroucheva, 2014). The fertilization of the ovules triggers the increase of auxin and gibberellin and a reduction of ABA (Obroucheva, 2014). It can be speculated that the higher levels of ABA directly measured in the receptacle of the artificially induced 'crumbly' green fruit and those indirectly inferred by the results of the microarray analysis in the green fruit of the mostly 'crumbly' plants might be responsible of an hormonal disequilibrium, with ABA levels too high compared to normal fruit development that might cause the formation of uneven fruit in raspberry.

The semi-quantitative analysis of target phytohormones, although it did not show any significant differences between the two phenotypes for three of the phytohormones identified (IAA, GA_4 and SA), it produced useful information about the role of these five plant hormones during the growth and development of normal fruit in both drupelet and receptacle at the green and red fruit stages. The results, with regards to the different parts analysed, showed that IAA was the only phytohormone whose amount, in the samples analysed, was significantly higher in the drupelet. On the other hand, all the other four compounds (i.e. ABA, GA_1, GA_4 and SA) were always statistically significantly higher in receptacle. The differences at the two stages (i.e. green and red fruit) showed that while for the two gibberellins (i.e. GA_1 and GA_4) the difference was not significant, IAA and SA were significantly higher at the green stage while the amount of ABA was significantly higher at the red stage.

The overview of the situation highlighted by the outcomes of the phytohormones analyses depicted a system where the five plant hormones detected/quantified participate in the regulation of fruit growth and development of both the drupelet and receptacle although with some differences. IAA would seem to function primarily in the drupelets while all the other phytohormones (i.e. ABA, GA_1, GA_4 and SA) would support the growth of the receptacle. Furthermore, the results of the analysis showed that while IAA and SA seemed to be employed more at green stage, the opposite was observed for ABA; no differences, for the development stages were observed for both the gibberellins.

The current models describing the hormonal regulation of fruit growth, considers two phases, one before and the other after fertilization. Auxins and gibberellins at relative low levels, counterbalanced by relative high levels of ABA and ethylene (Vriezen et al., 2008, Ozga and Reinecke, 2003, Obroucheva, 2014) maintain both flower and fruit primordia as inactive. As an effect of fertilization both ABA and ethylene decrease with a consequent increase in auxin levels that promote cell division and the synthesis or activation of gibberellins that stimulate cell enlargement (Serrani et al., 2008); this in both the ovary (fruit pericarp) and in the ovule (seed). Cytokinins also seem to play an important role during the early stage of fruit set, promoting cell division in both ovary and ovule. The interplay between cytokinins, auxins and gibberellins appear to be crucial for the early steps of fruit development as was demonstrated by parthenocarpic fruits in tomato as a consequence of exogenous application of these three phytohormones (Kumar et al., 2014).

The analysis, of the 'crumbly' induction experiments, highlighted the involvement of a fifth phytohormone, salicylic acid (SA). Like IAA, SA too was significantly higher at the green than red fruit stage indicating that it is required at a relatively larger amount at the beginning of fruit growth. Although SA is associated primarily with response to stresses and systemic acquired resistance (SAR), it is an important component of the signal transduction systems (Elwan and El-Hamahmy, 2009) and is an extensive modulator of IAA (LeClere et al., 2008).

In chapter 5 of this work (see section 5.4) a putative model of the hormonal regulation of fruit development was suggested. The scheme speculates on the interplay between IAA, GA_1, ABA and SA as a driver of the fruit growth with IAA modulating ABA and SA negative regulating IAA. The significantly higher levels of IAA in drupelets compared to the receptacle might suggest its important role in assisting cell division in the fertilized ovaries, and in modulating ABA in drupelets. While the significantly higher levels of SA and GA_1 might suggest, for SA the modulation of IAA that is in fact significant lower in the receptacle while the promotion of cell enlargement in the receptacle is performed by GA_1.

All these models suggested are reasonable, they require to be considered relative to the stage at which the samples were collected. Samples were picked approximately two weeks after anthesis and at that stage, extensive cell division in the forming fruit might have been already taken place; for instance, in cucumber the process ends within four

days after anthesis (Boonkorkaew et al., 2008). Thus, the conclusions from the data available, for what would concern the hormonal regulation of fruit development cannot be certain for the very early stage, simply because the process might have already occurred at the time of sample collection.

6.6 Genetic analysis of 'crumbly' fruit

6.6.1 Hormone regulation hypothesis

Although the plant material tested was different between the phytohormone study (i.e. drupelets and receptacles from green and red fruits collected from Glen Ample plants grown under controlled environment) and the genetic study (QTL and microarray analysis), conclusions can still be drawn. For the microarray experiment, samples were selected from individuals of a mapping population between Latham ('crumbly' donor parent) and Glen Moy (see chapter 3 for full details). These samples were from three different development stages (i.e. closed bud, open flower and green berry) and of the two different phenotypes (i.e. mostly and never 'crumbly').

The analysis of the microarray probes matching *Rubus idaeus* genes whose *Arabidopsis thaliana* orthologs have gene ontology annotation related to 'response to ABA' showed that ten probes were upregulated in the mostly 'crumbly' phenotypes. Such ontology terms indicate any kind of process triggered by ABA stimulus that cause change in the state or in the activity of a cell or an organism (i.e. movement, secretion, enzyme production, gene expression, etc.). Such finding, in a certain sense, validated the results of the phytohormones analysis giving further credit to them and confirming the potential role played by ABA, at the green stage in the development of 'crumbly' fruits.

The 'crumbly' microarray analysis showed an interesting result in terms of the hypothesis behind this work, of a hormonal crosstalk between receptacle and fertilised ovaries, that drives the synchronised growth of all the pollinated carpels. Many of the 'crumbly' microarray probes that were differentially expressed between the two phenotypes (i.e. mostly and never 'crumbly') matched *R. idaeus* genes whose *A. thaliana* ortholog had gene ontology terms related to 'hormonal signalling pathway', involving seven different class of hormones (i.e. auxins, gibberellins, SA, ABA, ethylene, brassinosteroids (Brs) and jasmonic acid). The hormonal signalling pathway refers to the activation/modulation of molecular signals by the binding of a hormone to a cell receptor resulting in modulation of a molecular process. It might be speculated that this signalling pathway is the means through which the molecular messages might be transmitted between receptacle and fertilised ovaries. The majority of the genes were downregulated in the mostly 'crumbly' phenotype suggesting that this downregulation could indicate a reduction in hormonal crosstalk that might contribute to the formation of 'crumbly' fruit. Currently the data

cannot provide an explicit conclusion because in the 'crumbly' microarray experiment, receptacle and drupelet were not separated in the analysis. The microarray still provides clues about the value of the putative receptacle centred hormonal growth regulation system.

The microarray experiment also produced some evidence about the involvement of cytokinins (CKs) in the model describing raspberry fruit growth at early stage (green berry). CKs regulate cell division and in concert with gibberellins even cell enlargement, but in contrast to IAA they are modulated by SA (Honda et al., 2017, Tiwari et al., 2013). Two probes with ontology terms related to CK biosynthesis/metabolism were upregulated at the green berry stage and of further note, these were mapped inside the two 'crumbly' QTLs identified on linkage group 3 (see chapter 4 for more details). This represent a validation of the results from the microarray experiment, giving further credit to the putative role played by these two genes in 'crumbly' fruit because they reside in an area of the genome that is significantly associated with 'crumbliness'. The up-regulation of these two genes would suggest the involvement of significant higher levels of CKs in the mostly 'crumbly' phenotype, a parallel can be drawn in the artificially induced 'crumbly' fruit samples and we can then speculate about the putative participation of CKs in the hypothetical hormonal model describing fruit growth regulation.

According to this hypothetical model, CKs, in the receptacle, might replace IAA and attenuate ABA levels, because CKs, in contrast to IAA, are not affected by salicylic acid. The putative contribution/involvement of CKs might help to explain the scenario found in the artificial 'crumbly' samples where a significant higher level of ABA was detected/quantified. In fact, these significant higher levels of ABA, in receptacle, could be the consequence of a hormonal disequilibrium caused by lower putative levels of CKs that cannot attenuate ABA levels. Such hormonal imbalance might be responsible for the development of 'crumbly' fruit.

It could be that in the artificially induced 'crumbly' fruits the damage which affects some of the pollination events is similar to the situation that occurs in unpollinated flowers (or in flowers damaged by means such as nectar load, bee foraging etc). In fact, without fertilization of the ovules, the triggering of crucial events such as synthesis and or activation of auxins, cytokinins and gibberellins would not take place (Obroucheva, 2014). The levels of ABA would not be attenuated by auxins and the flower gradually senescence (Ozga and Reinecke, 2003). In a certain way, it could be speculated that something similar might happen to the receptacle of the damaged flowers of the 'crumbly'

induction experiments where it is reasonable to assume a link with the obstructed interplay between ovaries and receptacle. Soon after fertilization, in the damaged receptacle, the activation of the aforementioned events (i.e. synthesis and/or activation of auxins, cytokinins and gibberellins) would take place only in a minimal way. This would explain the significant higher level of ABA in the receptacle and the senescence for the majority of the carpels (see Figure 6.1) that, although all hand pollinated, did not progress on compromising the normal growth and development of the fertilised ovaries.

Similar considerations might be taken for the red stage, the targeted phytohormones analysis (see chapter 5) showed an increase of ABA in both receptacle and drupelet compared to the green stage and this would be compatible with its role as a leading regulatory factor of fruit ripening in non-climacteric plants such as raspberry. The analysis registered a decrease of both IAA and SA in red fruit while gibberellins were again the only compounds whose difference between the two different stages were not significant. According to the results of the phytohormones analysis, in normal fruit growth the interplay between ABA, IAA and SA seems to play a crucial part in fruit ripening with SA that modulating IAA and then indirectly increasing ABA that in turn drives the ripening of the fruit.

Figure 6.1: Artificially induced 'crumbly' fruit at red and green stage.

Examples of artificially induced 'crumbly' fruit produced by pinning the tip and side of the receptacle with a needle soon after emasculation. Two days after the treatment when the stigmas of the flowers become receptive, all the carpels were hand pollinated to make sure all the carpels were pollinated. On the left a fruit at red stage and on the right one at green stage. These are typical of those found in nature.

From the 'crumbly' fruit perspective, the most important result of the target phytohormones analysis was the significantly lower levels of GA_1 in the drupelets of artificially induced 'crumbly' fruits compared to the control. As for the green stage, the hormonal model with five components (i.e. ABA, GA_1, IAA, GA_4 and SA) could not easily explain the data related to the significantly higher level of GA_1 in the drupelet samples of the artificially induced 'crumbly' samples. Once again, the contribution of cytokinins (CKs) was speculated as a putative model of hormonal regulation of fruit growth at late stage (i.e. red berry). The above-mentioned CKs would cause a hormonal disequilibrium with the result of reducing the levels of GA_1 in red drupelets and of delaying the fruit development. This might explain the peculiar growth and ripening of the artificially induced 'crumbly' samples. In fact, they tended to grow slower, with the longer times required by the drupelets to reach, in some cases (see Figure 6.1), the very extremely large size observed in this type of samples 'crumbly'. The significantly lower levels of GA_1 coupled with putative high levels of CKS might prolong the cell elongation step that in turn cause the abnormal growth of the drupelets compared to those of normal berry.

The phytohormones analysis of the artificially induced 'crumbly' fruits allowed the discovery of some physiological aspects related to the hormonal regulation of fruit growth and development and even to provide evidence that the damage of the flower, and in particular of the receptacle, can cause the formation of 'crumbly' fruit.

6.6.2 Genetic control of developmental processes

The analysis of the 'crumbly' microarray experiment highlighted many probes differently expressed between the two phenotypes (i.e. mostly and never 'crumbly' identified seven genes, that were implicated in aspects of the flower/fruit growth that could be linked with processes responsible for the formation of misshapen 'crumbly' like fruits (for details see chapter 3). Pollen formation and anomalies in the activity of pollen and specifically of pollen tube, were selected as key processes leading to the formation of 'crumbly' fruit. From the genetic point of view, the analysis of the microarray experiments suggested the importance of processes upstream of the fertilization event as key steps in the growth and development of 'crumbly' fruit.

Impairments in the formation of mature pollen could substantially impact the formation of normal shape fruit because this defective pollen could not complete its function. The A. thaliana gene (SLY1) encoding SLEEPY1, of the components of the SCFSLY-SUMO complex is responsible for the ubiquitination of DELLA proteins and then indirectly of their degradation through the 26S proteosome complex (Kim at al. 2015). In the mostly 'crumbly' phenotype, the downregulation of the probes matching the R. idaeus gene, ortholog of SLY1, might have affected the response to gibberellins, a class of hormones important for processes such as pollen development and maturation (Daviere and Achard 2013). The impaired modulation of DELLA might affect the response to gibberellins with consequent reduction of fertility due to production of flawed pollen (Kim et al. 2015).

The downregulation in the mostly 'crumbly' phenotype of HMG1 in pollen might have caused the formation of abnormal pollen due to lack of important component of the membranes (i.e. phytosterols) whose synthesis in the tapetum cell of the pollen depends primarily on HMGR1. The Rubus idaeus equivalent to the Arabidopsis thaliana HMG1 encodes the β-hydroxy-β-methylglutaryl-CoA reductase (HMGR1) a key enzyme, in the tapetum cells of the pollen, for the cytosolic mevalonate pathway (MVA) important for the biosynthesis of terpens, precursors of sterols and steroids (Suzuki et al. 2009).

The downregulation of a putative Niemann-Pick C1 protein involved in regulation of the sterol pathway, might cause defect in the gametogenesis (Feldamn et al. 2015) with consequent formation of poor viable pollen.

These three examples of alteration in the formation of viable pollen were clear examples of anomalous processes taking place in the mostly 'crumbly' phenotype that suggest clear impacts on 'crumbly' fruit because for every flawed pollen reaching the stigma of the carpel, a raspberry fruit would form lacking drupelet(s).

The analysis of the 'crumbly' microarray also identified four genes controlling the synthesis of important components of the cell wall that are indispensable during processes such as pollen germination and pollen tube elongation.

Firstly, GLUCURONOKINASE G, a member GHMP-kinase superfamily whose specific substrate is d-glucuronic acid. The gene is expressed in all plant tissues with a preference for pollen where it supplies the cell wall polymers indispensable for accompanying the expansion of the pollen tube (Pieslinger et al. 2009). Secondly, AT1G63180.1, encodes (UGE3), a UDP-D-glucose 4-epimerase, responsible for the conversion of UDP-galactose to UDP-glucose (UDP-glc). The enzyme is involved in pollen development (source TAIR

https://ui.arabidopsis.org/) and it appears to work alongside the Glucuronokinase G. UGE3 produces UDP-glc which is the precursor of UDP-glcA (UDP-glucuronate), the substrate of Glucuronokinase G. Both the enzymes, UGE3 and Glucuronokinase G, are expressed preferentially in pollen and contributes indirectly to supply the required components to the cell wall during the rapid expansion of the pollen tube (Pieslinger et al. 2009). The probes of the 'crumbly' microarray probes matching these two genes were upregulated in the mostly 'crumbly' phenotypes giving an indication that an inappropriate expression level of these genes might be deleterious for the synthesis of components of the cell wall especially in the fast growing pollen tube compromising its ability to expand and reach the ovule for the fertilization.

The pollen tube, the fast growing tissue of the plant is required to elongate and carry the gametocytes to the ovule, needs both an efficient supply of cell wall components to follow through the expansion of the tissue but also requires an efficient machinery to transfer all these compounds to the site of growth of the cells. The actin cytoskeleton constitutes the track through which the cell coordinates the organized movement of all the components that are required to accompany both membrane expansion and cell wall synthesis. In the mostly 'crumbly' phenotype. The third gene AT3G12110.1 encodes ACT11 a reproductive actin, actively expressed in pollen, required for the correct rearrangement of the actin cytoskeleton (Chang and Huang 2015) was significantly downregulated. In the mostly 'crumbly' plants, the low expression of the gene encoding ACT11 might be responsible for impairment in pollen germination and pollen tube elongation with consequent lack of fertilization that can cause the formation of misshapen fruits with lower number of drupelets such as the 'crumbly' ones.

The pollen tube needs to protrude inside the style of the carpel but once it reaches the ovule, inside the ovary, its elongation should terminate and the pollen tube should erupt to release the two sperm cells and accomplish the fertilization (Escobar-Restrepo et al. 2007). Plants have evolved a series of strategies for the successful reception of the pollen tube inside the embryo sac and in particular a recognition system between Receptor Like Kinase (RLK) expressed on the plasma-membrane of the synergid cells and putative ligand produced by the pollen tube. Without this recognition the pollen tube continues its growth inside the embryo sac, does not erupt releasing the sperm cells and thus fertilization does not take place (Escobar-Restrepo et al. 2007). In A. thaliana, FERONIA (FER) is the RLK responsible for the correct interaction between pollen tube and synergid cells (Haruta et al. 2018). In the mostly 'crumbly' plants the fourth gene, the *R. idaeus*

equivalent gene of FERONIA was downregulated suggesting that a potential cause of misshapen fruits with lower number of drupelets might be found in missed fertilization due to incorrect recognition between pollen tube and embryo sac.

6.6.3 Microarray probes residing within QTL

Some of the microarray probes, differently expressed between the two phenotypes (i.e. mostly and never 'crumbly'), were mapped inside the three 'crumbly' QTLs. The two original ones, previously identified by Graham et al., (2015), and a third one identified during this work. From the 'crumbly' fruit perspective and for all the aforementioned properties, these probes and more precisely the genes to which they were matched were important for validation of the microarray data but also as being potential markers in future for breeding.

These genes included flower anatomical and pollen functional disruptions that represent potential factors responsible for the formation of misshapen fruit.

Under-expression of the *Arabidopsis thaliana* gene (AT5G12210.1) which encodes for the β-subunit (RGTB1) of the RGT (Rab geranylgeranyl transferase), negatively affected polar growth of pollen tubes compromising the fertilization of the ovules and causing the formation of anomalous flowers with carpels much longer that the stamens (Hala et al. 2010; Gutkowska et al. 2015). The 'crumbly' microarray probe, matching the *R. idaeus* gene equivalent of the *A. thaliana* (AT5G12210.1), was down-regulated in the mostly 'crumbly' plants suggesting, for the reason discussed above, a potential important role in causing the formation of 'crumbly' like misshapen fruits. The microarray probe, matching the *R. idaeus*, ortholog of the *A. thaliana* (AT5G12210.1), maps inside the 'crumbly' QTL on linkage group three previously identified by Graham et al., (2015)

Aminoacyl-tRNA synthetases (AARSs) are essential enzymes catalysing the reaction responsible for the attachment of amino acids to their corresponding tRNAs. The *methionyl-tRNA synthetase* called OVA1 (ovule abortion 1) is an important AARSs acting in the mitochondria whose disruption cause ovule abortion as shown in *Arabidopsis thaliana* T-DNA insertion mutants (Berg et al., 2005). The microarray probe, matching the *R. idaeus* gene ortholog of the *A. thaliana* (OVA1), was upregulated in the mostly 'crumbly' plants, those displaying misshapen fruits; the probe maps inside the 'crumbly' QTL recently identified during this work on linkage group three (LG3). It

might be speculated that behind processes such as ovule fertilization and seed formation there is the participation of many regulators acting both up and downstream to OVA1. The over expression of OVA1 could affect the action of these other regulators causing a functional disequilibrium whose consequence could be ovule abortion and/or arrest of seed formation.

Abnormal flowers with reduced number of carpels and consequently of ovules might be regarded as potential factors causing 'crumbly' fruit. The less carpels per flower the lower the drupelets per fruit. A differentially expressed microarray probe that mapped within the 'crumbly' QTL, previously identified by Graham et al., (2015) on linkage group three, matched a *R. idaeus* gene whose *A. thaliana* ortholog (AT1G62360.1) encodes the KNOX transcription factor SHOOT MERISTEMLESS (STM) which is responsible for promoting carpel development and their associated placental tissues (Scofield et al., 2007). Experiments with RNA interference on the STM gene (STM-RNAi) showed the formation of anatomical defects at the level of the floral meristem (Scofield et al., 2007). The model for the role of STM in carpel development proposed by Scofield et al., (2007) involves another key regulator of flower organ development, the AGAMUS like protein (AG) and its repressors, with BELLRINGER being one of them. Although STM was upregulated in the crumbly fruit AG and BELLRINGER were not differently expressed between the two phenotypes (i.e. mostly and never 'crumbly'). This would suggest that an appropriate balance of all the three factors is necessary for normal flower development and that the upregulation of STM would be deleterious and potentially responsible for the formation of flowers with anatomical defects.

The analysis of the gene expression showed a clear genetic component behind the 'crumbly' fruit phenomenon. Genes involved in processes related to pollen formation, pollen tube elongation, pollen tube and embryo sac recognition as well as flower development, were differently expressed, with differences being statistically significant. These genes were deleterious for the processes they were involving in and in fact, for the mostly 'crumbly' plants, the impairment of all these processes might translate easily in higher chances to develop misshapen fruits. Lower level of viable pollen rather than anatomically flawed flowers or impaired pollen tube protrusion into the style or even compromised pollen tube and embryo sac recognition are all clear examples of anomalous phenomena whose main effect is the reduction of the fertilization. In extreme synthesis, the less flower fertilized, the lower drupelet developed in a berry and the higher chance for the plants to form 'crumbly' like misshapen fruits.

6.6.3.1 Microarray probes residing within 'crumbly' QTLs matching genes having G.O. ontology annotations others than flower development, hormones, pollen and transport

In chapter 3, the study of the function of those genes matched by the microarray probes significantly differently expressed, for what concerned the stage*phenotype interaction, focused only on those genes having gene ontology annotation related to: flower development, hormones, pollen and transport (see section 3.3 for full details about the choice of these four GO terms). This approach allowed to contain the number of genes to be analysed by focusing on specific aspects related mainly to early stages of fruit growth and development that could be linked with 'crumbly' fruit. The same strategy was followed for those microarray probes located within the three 'crumbly' QTLs (see chapter 4 for full details) and the results of this study was summarized before in section 6.6.3.

The analysis of the gene functions for all the genes, located within the three 'crumbly' QTLs but having GO annotations others than flower development, hormones, pollen and transport resulted in the identification of three new genes whose products could be linked with the formation of 'crumbly' fruit in red raspberry. None of these three genes was located within the QTL on LG1 while two were on the QTL on LG3 previously identified by Graham et al., (2015) and one was on the 'crumbly' QTL identified during this work on LG3. The *A. thaliana* genes orthologs of the *R. idaeus'* ones matched by the microarray probes were: AT4G26840.1, AT5G08080.3 and AT2G21540.1. The first gene AT4G26840.1 encodes SUMO 1 a post translational modification small ubiquitin-like modifier. In plants reproductive processes such as flower development, mega- and micro-gametogenesis, fertilization and embryogenesis are regulated by many factors with post translational protein modification within them (Liu *et al.,* 2017b). SUMO proteins are involved in important processes related to plant reproduction such as flowering time regulation, GA signalling pathway and gametophyte development (Liu *et al.,* 2017b). The microarray probe (CUST_6848_PI426541283) matching the *R. idaeus'* gene ortholog of AT4G26840.1 and encoding SUMO 1 was downregulated in the mostly 'crumbly' samples with differences being statistically significant (<0.001), for what concerned the stage*phenotype interaction, only at open flower stage (see Table A.6.1 in appendix). Such result would suggest that in the mostly 'crumbly' plants, the significant reduction in the expression levels of the gene encoding SUMO 1 could be responsible of anomalous

gametes development resulting in probable formation of abnormal pollen and/or ovules that could negatively affect the fertilization and contribute to the formation of uneven fruit with reduced number of drupelets.

The second gene AT5G08080.3, located in the 'crumbly' QTL identified by Graham et al., (2015), encodes the SYNTAXIN OF PLANTS 132 (SYP132), a SNARE (soluble N-ethylmaleimide-sensitive factor attachment protein receptors) protein that controls the ultimate fusion of a secretory vesicle with its target compartment (Li et al., 2019). SYP132, located in the plasma membrane (PM) of many tissues including the developing pollen, is expressed even in the early stages of pollen development and represent a putative candidate to mediate fusion processes during the mitotic divisions that are responsible for the formation of the tricellular pollen (Li et al., 2019). The microarray probe (CUST_2371_PI426541283) matching the *R. idaeus'* gene ortholog of AT5G08080.3 and encoding SYP132 was downregulated in the mostly 'crumbly' samples with differences being statistically significant (<0.001), for what concerned the stage*phenotype interaction, at both open flower and green berry stages (see Table A.6.1 in appendix). In the mostly 'crumbly' plants, the significant lower expression of the SYP132 gene could be responsible for anomalies in pollen development with consequent formation of putative unviable pollen. As a consequence, all the carpels pollinated with these pollen grains cannot develop in fleshy drupelets with the effect of contributing to the formation of misshapen fruit with lower number of drupelets.

The gene AT2G21540.1, located within the 'crumbly' QTL identified during this work on linkage group 3, encodes SEC14-LIKE 3 (SFH3) a Phosphatidylinositol/phosphatidylcholine transfer protein (PITP). This class of proteins is ubiquitous and its function consist in binding and exchanging one molecule of phosphatidyli-nositide (PtdIns) or phosphatidylcholine (PtdCho), to facilitate the transfer of these phospholipids among the different membrane compartments of eukaryotic cells (Mo et al., 2007). SFH3 may be involved in certain aspects of protein secretion and polarized membrane trafficking during the fertilization process (Mo et al., 2007); in fact processes requiring polarized membrane trafficking are: germination of pollen grains on the stigma, guidance of pollen tubes in the style and delivery of the sperm nuclei to the ovule (Edlund et al., 2004). The microarray probe (CUST_26373_PI426541283) matching the *R. idaeus'* gene ortholog of AT2G21540.1 and encoding SFH3 was downregulated in the mostly 'crumbly' samples in all the three stages tested (i.e. closed bud, open flower and green berry) but the differences, in the expression levels were

statistically significant (<0.001), for what concerned the stage*phenotype interaction, only at open flower stage (see Table A.6.1 in appendix). In the mostly 'crumbly' plants, during flowering, the significant lower expression of the SFH3 gene could be responsible for anomalies in one of the main processes of fertilization (i.e. pollen germination, pollen tube elongation and/or delivery of the sperm nuclei to the ovule) with the effect of compromising the fertilization and causing the of misshapen fruits with lower number of drupelets.

The analysis of the gene function, for these three genes located in a region of the *Rubus* genome strongly associated with 'crumbly' fruit showed again how the impairment in the early stage of the fertilization, due mainly to abnormal pollen and/or ovules, represent a putative key factor for the formation of 'crumbly' fruit in red raspberry.

6.7 Crumbly fruit and fruit ripening

Another important aspect, directly linked to the genetic aspects behind 'crumbly' fruit, was the association found between the three 'crumbly' QTLs and fruit ripening QTLs. This linkage was first found by Graham et al., (2015) and was later strengthened by re-analysing the fruit ripening phenotypic scores with the higher density markers, Genotype by Sequencing map (Hackett et al., 2018). Of particular interest was the association between the original 'crumbly' QTL on linkage 3 and three fruit ripening stages/traits (i.e. PCO_2, green fruit and fruit set). PCO_2 was the principal coordinate two, (principal coordinates (PCOs) was the approach used to combined scores on different dates to extract the principal sources of variation in ripening across the mapping population) (Graham et al., 2009) interpreted as a plant showing slow fruit development in May but with a rapid ripening during late Jun and/or early July (Graham et al., 2009). For the association with fruit set and green fruit, this was positive and could be expressed as the longer the fruit took to reach fruit set and green fruit stage the more likely to be 'crumbly' (Graham et al., 2015). These associations are very interesting and in future can be explored for genetic interpretations. A consideration might be raised for the potential predisposition of one of the two different class of raspberry varieties; floricane and primocane (see introduction for further details). The 'crumbly' versus fruit ripening QTLs association was performed on floricane plants (siblings of two floricane cultivars, Latham and Glen Moy) with biennial growing cycle and characterised by early fruiting compared to primocane that are even called autumn fruiting varieties. It cannot be excluded, a priori,

that differences in terms of predisposition towards 'crumbly' fruit exist between these two classes of raspberry varieties. This issue opens up the need for further research on this topic although a little study was conducted during this work to tackle this aspect. An ordinal regression analysis was conducted on the same list of 63 different genotypes (see Table A.4.1 in appendix) used in the validation pool (see section 4.3.6). The ordinal regression assumes that there is an underlying continuous trait of crumbliness, splits into groups for scoring (i.e. 'crumbly' and non 'crumbly') and looks at the effect of other variables, in this case the different genotypes, on the distribution across groups. The result of the analysis showed that the difference between floricane and primocane in terms of predisposition towards 'crumbliness' were not statistically significant.

6.8 Conclusion

This study into the bases of 'crumbly' fruit, produced many interesting findings that increased the insight about the biochemistry, the genetics and some physiological aspects behind the 'crumbly' fruit phenomenon.

The starting point was the hypothesis that a molecular mechanism, mediated by phytohormones through crosstalk and controlled by the receptacle, regulates and synchronises the growth of the many drupelets (i.e. on average minimum 60) forming the final fruit. This assumption raised a several questions to be addressed. The first about the phytohormones and their interplay, while the second question concerned the role of the receptacle as leading hub for the fruit growth process. Furthermore, the main question regarded the role of phytohormones in 'crumbly' fruit.

The analysis of the phytohormones was described in chapter 5 of this work. It proved that plant hormones are involved in the process of fruit development, a model of hormonal regulated fruit growth was proposed with the engagement of four measured (i.e. ABA, GAs, IAA and SA) and one speculated CKs. The model was coherent with previous findings on fruit growth (Vriezen et al., 2008, Ozga and Reinecke, 2003, Obroucheva, 2014, Serrani et al., 2008, Kumar et al., 2014), the only exception was SA, which previously has not been reported in soft fruit. SA is an important component of the signal transduction systems (Elwan and El-Hamahmy, 2009) and is an extensive modulator of IAA (LeClere et al., 2008). From the 'crumbly' perspective the most important result was

the significantly higher concentration of ABA found in the receptacle at the green fruit stage and the significantly lower concentration of GA_1 in the drupelet at the red fruit stage in the artificially induced 'crumbly' samples. The involvement of hormones in fruit growth was demonstrated with the role of specific growth regulators (i.e. ABA and GA_1) highlighted.

In the phase of the design of both experiment and analytical method, it was decided to include methylated phytohormones as the compounds to be targeted. Methylated plant hormones could be the ideal candidate for intra and intercellular transport of hormones, allowing plants to easily move hormones, short and long distance, in a specific site (i.e. tissue or organ) where they can conduct their regulating action (Qin et al., 2005); the hypothesis behind this question was that the signal would start from the receptacle and then moves to the surrounding drupelets to coordinate and synchronize the growth of the fruit. According to this hypothesis and with the ease of which methylated compounds are transported through the cells, they might expected to play an important role in regulating the fruit growth and high levels of them might be found in the receptacle where they could be synthesized. However, none of the methylated compounds were detected during the analysis of all the 120 samples. It is reasonable to think that these compounds act at extremely low concentration (fmol-pmol/g of fresh weight) and their detection would be arduous even treating them separately (Simura et al., 2018). Obviously the simultaneous measurement of many different compounds each with specific chemical properties requires a compromise, between all the different target analytes, in terms of both sample preparation and methodology of analysis that would impact those analyte naturally present at lower concentration and/or with features that make their analysis more challenging. After all, the adoption of analytical methods (i.e. immunological assay) for a single compound or a specific class of phytohormones (Du et al., 2012a), although conceptually valid becomes impractical in cases in which many different compounds are under analysis simultaneously, as was the case here.

On the basis of the above considerations, it was evident that no conclusions could be drawn from the data of the LC-MS phytohormones analysis on its own. Support was needed from the gene expression analysis (see chapter 3 for full details).Some clues about this hypothetical hormonal crosstalk were determined from the microarray probes, differentially expressed between the two phenotypes (i.e. mostly and never 'crumbly'). Nine different classes of hormones (i.e. auxins, gibberellins, SA, ABA, ethylene, brassinosteroids (Brs) and jasmonic acid) were represented suggesting the participation of multiple different factors to this transfer of molecular messages. The majority of these

probes were downregulated suggesting that the reduction of this hormonal signalling might be associated with 'crumbliness' supporting this theory.

The same considerations for the methylated phytohormones applies to cytokinins (CKs) and more precisely to *trans*-zeatin, the only target CK analysed in this work. It was not detected in the samples but its involvement in the hormonal fruit growth model was speculated because for the mostly 'crumbly' plants, both the significantly higher level of ABA in receptacle samples at green stage and significant low level of GA_1 in drupelet at red fruit stage. The decision to select the *trans*-zeatin as an additional hormone to complete the model was supported by both literature data, confirming the important role play by this growth regulator in fruit development (Kumar et al., 2014) and the results of the 'crumbly' microarray experiment. In the latter, two probes significantly differentially expressed between the two phenotypes (i.e. mostly and never 'crumbly') matched two *R. idaeus* genes whose *A. thaliana* orthologs have gene ontology terms related to CKs biosynthesis/metabolism. The two probes were significantly upregulated in the mostly crumbly plants at the green berry stage suggesting that higher level of CKs might be needed during the development of 'crumbly' fruit at the green stage. Moreover, to give further credit to these two genes, with respect to 'crumbly' fruit, was the position along the linkage map. They mapped inside the two 'crumbly' QTLs on linkage group 3, strongly associated with 'crumbly' fruit.

The analysis of gene expression clearly demonstrated that 'crumbly' fruit had a genetic base. Pollen was identified as a key factor and the genes involved in its formation and functioning seemed to play a crucial role in the formation of 'crumbly' like misshapen fruits. The mostly 'crumbly' fruit were shown to have impairment in the formation of viable pollen or anomalous expression of genes controlling the biosynthesis of important cell wall components due to the lack of structural component and/or impairment of the machinery (i.e. actin cytoskeleton) indispensable to deliver these compounds (i.e. sterols above all) and sustain a very demanding process such as the protrusion of the pollen tube. Defects in the molecular component indispensable for the recognition between pollen tube and embryo sac were other examples of impaired processes, caused by abnormal expression, always in the mostly 'crumbly' plants of the gene involved in these events. Lastly it was clear that the final effect of all this impaired process resulted in no fertilization of the ovules which translates into berries with a reduced number of drupelets, that for this reason might give raise to 'crumbly' like misshapen fruits.

This work focused on the study of Malformed Fruit Disorder (MFD) in raspberry by means of studying the genetics/gene expression and hormone profiles.

6.9 Future work

This work identified other areas for study into the triggers of crumbly fruit. A new and interesting potential trigger of 'crumbly' fruit was observed related to flower nectar. This was not directly investigated, though excess nectar was observed as described below, during this work and a few considerations about its impact on 'crumbly' fruit in this discussion may give a better view of potential strategies that might be considered by the raspberry industry to try to tackle the problem.

Some observations on raspberry plants grown under glasshouse conditions, with the assistance of commercial beehives to increase pollination efficiency, highlighted an interesting phenomenon. Excess floral nectar not collected by bees was observed on the carpels. These became impregnated with the sticky substance and could not be pollinated. Given that the part of the flower affected by this phenomenon was huge, the result was the formation of misshapen, 'crumbly' like fruit. Floral nectar is a sugary secretion produced by special organ (i.e. nectaries) usually located inside the flower next to the reproductive organs (Nepi et al., 2018; Roy et al., 2017). Plants produce floral nectar to attract insect pollinators and in exchange of their service (pollination) plants provide the insects with a ready-to-use energy source (Nepi et al., 2018). Bees are the major insect pollinator used for commercial purposes and two main genera dominate the market, bumble bee (*Bombus ssp*) and honeybee (*Apis mellifera*). Bumble bees are recommended for greenhouse-grown plants because they adapt easier to closed spaces (Andrikopoulos and Cane, 2018). Bumble bees are twice as likely to forage for pollen than nectar compared to honey bees (Andrikopoulos and Cane, 2018) and to further encourage bumble bees to focus exclusively in collecting pollen, commercial hives are equipped with small tanks containing artificial nectar to supply anytime the bees with this substance so they do not need to collect nectar and can concentrate on foraging only pollen. This particular design of the hives was specifically studied to increase pollination efficiency. Honeybees when foraging, primarily, look for nectar. Further investigation would be required on this topic but probably a better management of the pollinators, for instance

combining together bumble bees and honeybees to exploit the at the maximum their different foraging behaviour could help industry not only to increase pollination efficiency but potentially contribute to the reduction of misshapen fruit due to excess of nectar impregnating flower carpels.

Another trigger observed during this work was temperature, where replicate clonal plants developed the crumbly phenomenon to different extents under different conditions. Data was collected on a range of environmental parameters and this needs to be investigated to identify what the triggers are.

A third observation made during this thesis was the lack of evidence for the 'crumbly' phenomena produced by mutation during the clonal propagation of plants. Four varieties were propagated under identical controlled conditions and planted in a randomized field experiment. To date no evidence of the crumbly phenotype has been observed but studies need to continue.

Another area for study would be agrochemical fine-tuned treatments with ABA inhibitors at the emergence of the symptoms to investigate their role as a potential control strategy.

The identification of robust 'crumbly' QTLs and their corresponding associated markers supported by gene expression data allow the possibility of the identification of markers for breeding using marker assisted selection of potentially 'crumbly' free varieties. In this work, a first attempt at the selection of 'crumbly' markers did not identify any associations in the limited number of polymorphisms investigated but further work here can select larger genome regions and SNPs and continue to look for associations. SSR markers identified have yet to be tested.

The 'crumbly' phenomenon is very complex, many variables are involved not just genetic and physiology, but with the environment playing a very important role. For its control, it is important that industry adopt a more holistic approach that consider all the possible aspects, considered in this work, that might affect fruit growth and development. Breeding of new varieties 'crumbly' free is the best approach in the medium-long term but before the release of these new cultivars, simpler and easy to practice measurement might be adopted by industry. A Genome Wide Association Study (GWAS) with a big population to get enough recombination could be performed to look for markers associated with the 'crumbly' phenotype that could be used to speed up the breeding programmes and then the selection of new varieties 'crumbly' free. Another relatively easily practical measure could be, for instance, the design of trials to investigate the

importance of a better bee pollinators management (i.e. combining honeybees and bumble bees to exploit their different foraging behaviour) that in the short term could help to tackle the problem and help the raspberry industry to reduce yield losses.

Reference

ABE, M., KOBAYASHI, Y., YAMAMOTO, S., DAIMON, Y., YAMAGUCHI, A., IKEDA, Y., ICHINOKI, H., NOTAGUCHI, M., GOTO, K. & ARAKI, T. 2005. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science,* 309**,** 1052-1056.

ALICE, L. A. & CAMPBELL, C. S. 1999. Phylogeny of Rubus (Rosaceae) based on nuclear ribosomal DNA internal transcribed spacer region sequences. *American Journal of Botany,* 86**,** 81-97.

ANDRIKOPOULOS, C. J. & CANE, J. H. 2018. Comparative Pollination Efficacies of Five Bee Species on Raspberry. *J Econ Entomol,* 111**,** 2513-2519.

ANTONIUS-KLEMOLA, K. 1999. Molecular markers in Rubus (Rosaceae) research and breeding. *The Journal of Horticultural Science and Biotechnology,* 74**,** 149-160.

ANTTONEN, M. J. & KARJALAINEN, R. O. 2005. Environmental and genetic variation of phenolic compounds in red raspberry. *Journal of Food Composition and Analysis,* 18**,** 759-769.

ARMBRUSTER, D. A. & PRY, T. 2008. Limit of blank, limit of detection and limit of quantitation. *Clin Biochem Rev,* 29 Suppl 1**,** S49-52.

BASSIL, N. V., DAVIS, T. M., ZHANG, H., FICKLIN, S., MITTMANN, M., WEBSTER, T., MAHONEY, L., WOOD, D., ALPERIN, E. S., ROSYARA, U. R., KOEHORST-VANC PUTTEN, H., MONFORT, A., SARGENT, D. J., AMAYA, I., DENOYES, B., BIANCO, L., VAN DIJK, T., PIRANI, A., IEZZONI, A., MAIN, D., PEACE, C., YANG, Y., WHITAKER, V., VERMA, S., BELLON, L., BREW, F., HERRERA, R. & VAN DE WEG, E. 2015. Development and preliminary evaluation of a 90 K Axiom(R) SNP array for the allo-octoploid cultivated strawberry Fragaria x ananassa. *BMC Genomics,* 16**,** 155.

BEEKWILDER, J., JONKER, H., MEESTERS, P., HALL, R. D., VAN DER MEER, I. M. & DE VOS, C. H. R. 2005. Antioxidants in raspberry: On-line analysis links antioxidant activity to a diversity of individual metabolites. *Journal of Agricultural and Food Chemistry,* 53**,** 3313-3320.

BEISSINGER, T. M., HIRSCH, C. N., SEKHON, R. S., FOERSTER, J. M., JOHNSON, J. M., MUTTONI, G., VAILLANCOURT, B., BUELL, C. R., KAEPPLER, S. M.

& DE LEON, N. 2013. Marker Density and Read Depth for Genotyping Populations Using Genotyping-by-Sequencing. *Genetics,* 193**,** 1073-1081.

BERG, M., ROGERS, R., MURALLA, R. & MEINKE, D. 2005. Requirement of aminoacyl-tRNA synthetases for gametogenesis and embryo development in Arabidopsis. *Plant Journal,* 44**,** 866-878.

BIAŁA, W. & JASIŃSKI, M. 2018. Phenylpropanoid Case – It Is Transport That Matters. *Front. Plant Sci.,* 9**,** 8.

BINET, R., FERNANDEZ, R. E., FISHER, D. J. & MAURELLI, A. T. 2011. Identification and characterization of the Chlamydia trachomatis L2 S-adenosylmethionine transporter. *mBio,* 2**,** e00051-11.

BOONKORKAEW, P., HIKOSAKA, S. & SUGIYAMA, N. 2008. Effect of pollination on cell division, cell enlargement, and endogenous hormones in fruit development in a gynoecious cucumber. *Scientia Horticulturae,* 116**,** 1-7.

BOTTCHER, C., BOSS, P. K. & DAVIES, C. 2011. Acyl substrate preferences of an IAA-amido synthetase account for variations in grape (Vitis vinifera L.) berry ripening caused by different auxinic compounds indicating the importance of auxin conjugation in plant development. *Journal of Experimental Botany,* 62**,** 4267-4280.

BURCH-SMITH, T. M., CUI, Y. & ZAMBRYSKI, P. C. 2012. Reduced levels of class 1 reversibly glycosylated polypeptide increase intercellular transport via plasmodesmata. *Plant Signal Behav,* 7**,** 62-7.

BUSHAKRA, J. M., BRYANT, D. W., DOSSETT, M., VINING, K. J., VANBUREN, R., GILMORE, B. S., LEE, J., MOCKLER, T. C., FINN, C. E. & BASSIL, N. V. 2015. A genetic linkage map of black raspberry (Rubus occidentalis) and the mapping of Ag (4) conferring resistance to the aphid Amphorophora agathonica. *Theoretical and Applied Genetics,* 128**,** 1631-1646.

BUSHAKRA, J. M., STEPHENS, M. J., ATMADJAJA, A. N., LEWERS, K. S., SYMONDS, V. V., UDALL, J. A., CHAGNE, D., BUCK, E. J. & GARDINER, S. E. 2012. Construction of black (Rubus occidentalis) and red (R. idaeus) raspberry linkage maps and their comparison to the genomes of strawberry, apple, and peach. *Theoretical and Applied Genetics,* 125**,** 311-327.

CANE, J. H. 2005. Pollination potential of the bee Osmia aglaia for cultivated red raspberries and blackberries (Rubus: Rosaceae). *HortScience,* 40**,** 1705 - 1708.

CANET, J. V., DOBON, A. & TORNERO, P. 2012. Non-recognition-of-BTH4, an Arabidopsis mediator subunit homolog, is necessary for development and response to salicylic acid. *Plant Cell,* 24**,** 4220-35.

CAO, K., CUI, L. R., ZHOU, X. T., YE, L., ZOU, Z. R. & DENG, S. L. 2016a. Four Tomato FLOWERING LOCUS T-Like Proteins Act Antagonistically to Regulate Floral Initiation. *Frontiers in Plant Science,* 6.

CAO, Z. Y., SUN, L. H., MOU, R. X., ZHANG, L. P., LIN, X. Y., ZHU, Z. W. & CHEN, M. X. 2016b. Profiling of phytohormones and their major metabolites in rice using binary solid-phase extraction and liquid chromatography-triple quadrupole mass spectrometry. *Journal of Chromatography A,* 1451**,** 67-74.

CASTILLO, A., DORADO, G., FEUILLET, C., SOURDILLE, P. & HERNANDEZ, P. 2010. Genetic structure and ecogeographical adaptation in wild barley (Hordeum chilense Roemer et Schultes) as revealed by microsatellite markers. *BMC Plant Biology,* 10.

CASTRO, P., STAFNE, E. T., CLARK, J. R. & LEWERS, K. S. 2013. Genetic map of the primocane-fruiting and thornless traits of tetraploid blackberry. *Theoretical and Applied Genetics,* 126**,** 2521-2532.

CHANG, M. & HUANG, S. J. 2015. Arabidopsis ACT11 modifies actin turnover to promote pollen germination and maintain the normal rate of tube growth. *Plant Journal,* 83**,** 515-527.

CHEN, Y., ZOU, T. & MCCORMICK, S. 2016. S-Adenosylmethionine Synthetase 3 Is Important for Pollen Tube Growth. *Plant Physiology,* 172**,** 244-253.

COLLARD, B. C. Y., JAHUFER, M. Z. Z., BROUWER, J. B. & PANG, E. C. K. 2005. An introduction to markers, quantitative trait loci (QTL) mapping and marker-assisted selection for crop improvement: The basic concepts. *Euphytica,* 142**,** 169-196.

DAVEY, J. E. & VANSTADEN, J. 1978. Cytokinin Activity in Lupinus-Albus .2. Distribution in Fruiting Plants. *Physiologia Plantarum,* 43**,** 82-86.

DAVIERE, J. M. & ACHARD, P. 2013. Gibberellin signaling in plants. *Development,* 140**,** 1147-1151.

DI MARZO, M., HERRERA-UBALDO, H., CAPORALI, E., MENDES, M. A., DE FOLTER, S. & BCOLOMBO, L. 2020. SEEDSTICK Controls Arabidopsis Fruit Size by Regulating Cytokinin Levels and FRUITFULL. *Cell Reports,* 30**,** 2846-2857.

DOBSON, P., GRAHAM, J., STEWART, D., BRENNAN, R., HACKETT, C. A. & MCDOUGALL, G. J. 2012. Over-seasons Analysis of Quantitative Trait Loci Affecting Phenolic Content and Antioxidant Capacity in Raspberry. *Journal of Agricultural and Food Chemistry,* 60**,** 5360-5366.

DOLAN, A. 2018. Pathogen Testing Requirements for Raspberry Material Entering the EU Certification Scheme. *In:* GRAHAM, J. & BRENNAN, R. (eds.) *Raspberry: Breeding, Challenges and Advances.* Springer International Publishing, pp 83-88.

DOSSETT, M., LEE, J. & FINN, C. E. 2008. Inheritance of phenological, vegetative, and fruit chemistry traits in Black raspherry. *Journal of the American Society for Horticultural Science,* 133**,** 408-417.

DU, F. Y., RUAN, G. H. & LIU, H. W. 2012. Analytical methods for tracing plant hormones. *Analytical and Bioanalytical Chemistry,* 403**,** 55-74.

EBEL, C., MARICONTI, L. & GRUISSEM, W. 2004. Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. *Nature,* 429**,** 776-780.

EDLUND, A. F., SWANSON, R. & PREUSS D. 2004. Pollen and stigma structure and function: the role of diversity in pollination. Plant Cell, 16, 84-97.

ELSHIRE, R. J., GLAUBITZ, J. C., SUN, Q., POLAND, J. A., KAWAMOTO, K., BUCKLER, E. S. & MITCHELL, S. E. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *Plos One,* 6.

ELWAN, M. W. M. & EL-HAMAHMY, M. A. M. 2009. Improved productivity and quality associated with salicylic acid application in greenhouse pepper. *Scientia Horticulturae,* 122**,** 521-526.

EMENECKER, R. J. & STRADER, L. C. 2020. Auxin-abscisic acid interactions in plant growth and development. Biomolecules, 10, 281.

ESCOBAR-RESTREPO, J. M., HUCK, N., KESSLER, S., GAGLIARDINI, V., GHEYSELINCK, J., YANG, W. C. & GROSSNIKLAUS, U. 2007. The FERONIA receptor-like kinase mediates male-female interactions during pollen tube reception. *Science,* 317**,** 656-60.

FELDMAN, M. J., POIRIER, B. C. & LANGE, B. M. 2015. Misexpression of the Niemann-Pick disease type C1 (NPC1)-like protein in Arabidopsis causes sphingolipid accumulation and reproductive defects. *Planta,* 242**,** 921-933.

FINN, C. C. & HANCOCK, J. F. 2008. Raspberries. *In:* HANCOCK J. F. (ed) Temperate fruit crop breeding: germplam to genomics. Springer Netherlands, pp 359-392

FORNARA, F., DE MONTAIGU, A. & COUPLAND, G. 2010. SnapShot: Control of Flowering in Arabidopsis. *Cell,* 141.

FOSTER, T. M., BASSIL, N. V., DOSSETT, M., WORTHINGTON, M. L. & GRAHAM, J. 2019. Genetic and genomic resources for Rubus breeding: a roadmap for the future. *Horticulture Research,* 6.

GARCIA-MAZCORRO, J. F., PEDRESCHI, R., CHEW, B., DOWD, S. E., KAWAS, J. R. & NORATTO, G. 2018. Dietary Supplementation with Raspberry Extracts Modifies the Fecal Microbiota in Obese Diabetic db/db Mice. *J Microbiol Biotechnol,* 28**,** 1247-1259.

GARDENER, M. C. & GILLMAN, M. P. 2002. The taste of nectar - a neglected area of pollination ecology. *Oikos,* 98**,** 552-557.

GENDRE, D., MCFARLANE, H. E., JOHNSON, E., MOUILLE, G., SJODIN, A., OH, J., LEVESQUE-TREMBLAY, G., WATANABE, Y., SAMUELS, L. & BHALERAO, R. P. 2013. Trans-Golgi network localized ECHIDNA/Ypt interacting protein complex is required for the secretion of cell wall polysaccharides in Arabidopsis. *Plant Cell,* 25**,** 2633-46.

GIVEN, N. K., VENIS, M. A. & GRIERSON, D. 1988. Hormonal-Regulation of Ripening in the Strawberry, a Non-Climacteric Fruit. *Planta,* 174**,** 402-406.

GOETZ, M., HOOPER, L. C., JOHNSON, S. D., RODRIGUES, J. C. M., VIVIAN-SMITH, A. & KOLTUNOW, A. M. 2007. Expression of aberrant forms of AUXIN RESPONSE FACTOR8 stimulates parthenocarpy in Arabidopsis and tomato. *Plant Physiology,* 145**,** 351-366.

GRAHAM, J. & BRENNAN, R. 2018. Introduction to the *Rubus* Genus. *In:* GRAHAM, J. & BRENNAN, R. (eds.) *Raspberry: Breeding, Challenges and Advances.* 1 ed. Switzerland: Springer International Publishing, pp 1-16.

GRAHAM, J., SMITH, K., WOODHEAD, M. & RUSSELL, J. 2002. Development and use of simple sequence repeat SSR markers in Rubus species *Molecular Ecology Resources,* 2**,** 250-252.

GRAHAM, J., HACKETT, C. A., SMITH, K., WOODHEAD, M., HEIN, I. & MCCALLUM, S. 2009. Mapping QTLs for developmental traits in raspberry from bud break to ripe fruit. *Theoretical and Applied Genetics,* 118**,** 1143-1155.

GRAHAM, J., HACKETT, C. A., SMITH, K., WOODHEAD, M., MACKENZIE, K., TIERNEY, I., COOKE, D., BAYER, M. & JENNINGS, N. 2011. Towards an understanding of the nature of resistance to Phytophthora root rot in red raspberry. *Theoretical and Applied Genetics,* 123**,** 585-601.

GRAHAM, J. & SIMPSON, C. 2018. Developmental Transitions to Fruiting in Red Raspberry *In:* HYTÖNEN, T., GRAHAM, J. & HARRISON, R. (eds.) *The Genomes of Rosaceous Berries and Their Wild Relatives.* Springer, Cham.

GRAHAM, J., SMITH, K., MACKENZIE, K., JORGENSON, L., HACKETT, C. & POWELL, W. 2004. The construction of a genetic linkage map of red raspberry (Rubus idaeus subsp. idaeus) based on AFLPs, genomic-SSR and EST-SSR markers. *Theor Appl Genet,* 109**,** 740-9.

GRAHAM, J., SMITH, K., MCCALLUM, S., HEDLEY, P. E., CULLEN, D. W., DOLAN, A., MILNE, L., MCNICOL, J. W. & HACKETT, C. A. 2015. Towards an understanding of the control of 'crumbly' fruit in red raspberry. *Springerplus,* 4.

GRAHAM, J., SMITH, K., TIERNEY, I., MACKENZIE, K. & HACKETT, C. A. 2006. Mapping gene H controlling cane pubescence in raspberry and its association with resistance to cane botrytis and spur blight, rust and cane spot. *Theoretical and Applied Genetics,* 112**,** 818-831.

GUPTA, P. K., VARSHNEY, R. K., SHARMA, P. C. & RAMESH, B. 1999. Molecular markers and their applications in wheat breeding. *Plant Breeding,* 118**,** 369-390.

GUTKOWSKA, M., WNUK, M., NOWAKOWSKA, J., LICHOCKA, M., STRONKOWSKI, M. M. & SWIEZEWSKA, E. 2015. Rab geranylgeranyl transferase beta subunit is essential for male fertility and tip growth in Arabidopsis. *Journal of Experimental Botany,* 66**,** 213-224.

HACKETT, C. A. 2002. Statistical methods for QTL mapping in cereals. *Plant Mol Biol,* 48**,** 585-99.

HACKETT, C. A., MCLEAN, K. & BRYAN, G. J. 2013. Linkage Analysis and QTL Mapping Using SNP Dosage Data in a Tetraploid Potato Mapping Population. *Plos One,* 8.

HACKETT, C. A., MILNE, L., SMITH, K., HEDLEY, P., MORRIS, J., SIMPSON, C. G., PREEDY, K. & GRAHAM, J. 2018. Enhancement of Glen Moy x Latham raspberry linkage map using GbS to further understand control of developmental processes leading to fruit ripening. *Bmc Genetics,* 19.

HALA, M., SOUKUPOVA, H., SYNEK, L. & ZARSKY, V. 2010. Arabidopsis RAB geranylgeranyl transferase beta-subunit mutant is constitutively photomorphogenic, and has shoot growth and gravitropic defects. *Plant Journal,* 62**,** 615-627.

HANCOCK, R. D., PETRIDIS, A. & MCDOUGALL, G. J. 2018. Raspberry fruit chemistry in relation to fruit quality and human nutrition. *In:* GRAHAM, J. & BRENNAN, R. (eds.) *Raspberry: breeding, challenges and advances.* 1 ed. Springer International Publishing, pp 89-119

HARPSTER, M. H., BRUMMELL, D. A. & DUNSMUIR, P. 1998. Expression analysis of a ripening-specific, auxin-repressed endo-1, 4-beta-glucanase gene in strawberry. *Plant Physiol,* 118**,** 1307-16.

HARUTA, M., GADDAMEEDI, V., BURCH, H., FERNANDEZ, D. & SUSSMAN, M. R. 2018a. Comparison of the effects of a kinase-dead mutation of FERONIA on ovule fertilization and root growth of Arabidopsis. *Febs Letters,* 592**,** 2395-2402.

HARUTA, M., GADDAMEEDI, V., BURCH, H., FERNANDEZ, D. & SUSSMAN, M. R. 2018b. Comparison of the effects of a kinase-dead mutation of FERONIA on ovule fertilization and root growth of Arabidopsis. *FEBS Lett,* 592**,** 2395-2402.

HE, J. F., ZHAO, X. Q., LAROCHE, A., LU, Z. X., LIU, H. K. & LI, Z. Q. 2014. Genotyping-by-sequencing (GBS), an ultimate marker-assisted selection (MAS) tool to accelerate plant breeding. *Frontiers in Plant Science,* 5.

HEIDE, O. M. & SONSTEBY, A. 2011. Physiology of flowering and dormancy regulation in annual- and biennial-fruiting red raspberry (Rubus idaeus L.) - a review. *Journal of Horticultural Science & Biotechnology,* 86**,** 433-442.

HOKANSON, S. C., LAMBOY, W. F., SZEWC-MCFADDEN, A. K. & MCFERSON, J. R. 2001. Microsatellite (SSR) variation in a collection of Malus (apple) species and hybrids. *EUPHYTICA,* 118**,** 281-294.

HONDA, I., MATSUNAGA, H., KIKUCHI, K., MATUO, S., FUKUDA, M. & IMANISHI, S. 2017. Involvement of Cytokinins, 3-Indoleacetic Acid, and Gibberellins in Early Fruit Growth in Pepper (Capsicum annuum L.). *Horticulture Journal,* 86**,** 52-60.

HOSHI, A., I., YOSHIKO, K., SHIGEYUKI, O., KENRO, & SHIGETOU, N. 2007. Host-parasite interaction of phytoplasmas from a molecular biological perspective. *Bullettin of Insectology*, 60, 105-107.

HUAN, C., JIANG, L., AN, X. J., KANG, R. Y., YU, M. L., MA, R. J. & YU, Z. F. 2016. Potential role of glutathione peroxidase gene family in peach fruit ripening under combined postharvest treatment with heat and 1-MCP. *Postharvest Biology and Technology,* 111**,** 175-184.

HUANG, H., GAO, H., LIU, B., QI, T., TONG, J., XIAO, L., XIE, D. & SONG, S. 2017. Arabidopsis MYB24 Regulates Jasmonate-Mediated Stamen Development. *Front Plant Sci,* 8**,** 1525.

JANG, Y. H., PARK, H., LEE, K. C., THU, M. P., KIM, S., SUH, M. C., KANG, H. & KIM, J. 2014. A homolog of splicing factor SF1 is essential for development and is involved in the alternative splicing of pre-mRNA in Arabidopsis thaliana. *the plant journal,* 78**,** 591-603.

JENNINGS, D. L. 1967. Balanced Lethals and Polymorphism in Rubus Idaeus. *Heredity,* 22**,** 465-479.

JENNINGS, D. L. 1967b. Observations on some instances of partial sterility in red raspberry cultivars. *Horticultural Research,* 7**,** 116–122.

JENNINGS, D. L. 1988. *Raspberries and Blackberries: Their Breeding, Diseases and Growth,* London, Academic Press Limited.

JENNINGS, S. N. 2018. Advances in Rubus Breeding. *In:* GRAHAM, J. & BRENNAN, R. (eds.) *Raspberry: Breeding, Challenges and Advances.* Springer International Publishing.

JENNINGS, S. N. & BRENNAN, R. M. 2002. Improvement of raspberry cultivars in Scotland. *Proceedings of the Eighth International Rubus and Ribes Symposium, Vols 1 and 2***,** 179-183.

JIANG, T., X., Z., WANG, X. & ZHANG, D. 2011. Arabidopsis 3-Ketoacyl-CoA Thiolase-2 (KAT2), an Enzyme of Fatty Acid β-Oxidation, is Involved in ABA Signal Transduction. *Plant and Cell Physiology,* 52**,** 528–538.

JONES, C. S., IANNETTA, P. P. M., WOODHEAD, M., DAVIES, H., MCNICOL, R. J. & TAYLOR, M. A. 1997. The isolation of RNA from raspberry (Rubus idaeus) fruit. Molecular Biotechnology, 8, 219-221.

KANG, B. H., RANCOUR, D. M. & BEDNAREK, S. Y. 2003. The dynamin-like protein ADL1C is essential for plasma membrane maintenance during pollen maturation. *Plant Journal,* 35**,** 1-15.

KASSIM, A., POETTE, J., PATERSON, A., ZAIT, D., MCCALLUM, S., WOODHEAD, M., SMITH, K., HACKETT, C. & GRAHAM, J. 2009. Environmental and seasonal influences on red raspberry anthocyanin antioxidant contents and identification of quantitative traits loci (QTL). *Molecular Nutrition & Food Research,* 53**,** 625-634.

KELLOGG, A. A., BRANAMAN, T. J., JONES, N. M., LITTLE, C. Z. & SWANSON, J. D. 2011. Morphological studies of developing Rubus prickles suggest that they are modified glandular trichomes. *Botany-Botanique,* 89**,** 217-226.

KIM, S. I., PARK, B. S., KIM, D. Y., YEU, S. Y., SONG, S. I., SONG, J. T. & SEO, H. S. 2015. E3 SUMO ligase AtSIZ1 positively regulates SLY1-mediated GA signalling and plant development. *Biochemical Journal,* 469**,** 299-314.

KOEBNER, R. M. & SUMMERS, R. W. 2003. 21st century wheat breeding: plot selection or plate detection? *Trends Biotechnol,* 21**,** 59-63.

KONDO, S., SUNGCOME, K., SETHA, S. & HIRAI, N. 2004. ABA catabolism during development and storage in mangoes: Influence of jasmonates. *Journal of Horticultural Science & Biotechnology,* 79**,** 891-896.

KOWALCZYK, M. & SANDBERG, G. 2001. Quantitative analysis of indole-3-acetic acid metabolites in Arabidopsis. *Plant Physiology,* 127**,** 1845-1853.

KSHATRIYA, D., LI, X. Y., GIUNTA, G. M., YUAN, B., ZHAO, D. Y., SIMON, J. E., WU, Q. L. & BELLO, N. T. 2019. Phenolic-enriched raspberry fruit extract (Rubus idaeus) resulted in lower weight gain, increased ambulatory activity, and elevated hepatic lipoprotein lipase and heme oxygenase-1 expression in male mice fed a high-fat diet. *Nutrition Research,* 68**,** 19-33.

KUMAR, R., KHURANA, A. & SHARMA, A. K. 2014. Role of plant hormones and their interplay in development and ripening of fleshy fruits. *Journal of Experimental Botany,* 65**,** 4561-4575.

LECLERE, S., SCHMELZ, E. A. & CHOUREY, P. S. 2008. Cell wall invertase-deficient miniature1 kernels have altered phytohormone levels. *Phytochemistry,* 69**,** 692-699.

LI, B., LI, Y., LIU, F., TAN, X., RUI, Q., TONG, Y. et al., 2019. Overexpressed Tomosyn binds syntaxins and blocks secretion during pollen development. *Plant Physiology*, 181, 1114-1126.

LINCK, H. & REINEKE, A. 2019. Rubus stunt: a review of an important phytoplasma disease in Rubus spp. *Journal of Plant Diseases and Protection,* 126**,** 393-399.

LINCOLN, C., BRITTON, J. H. & ESTELLE, M. 1990. Growth and development of the axr1 mutants of Arabidopsis. *Plant Cell,* 2**,** 1071-80.

LIU, Y. & LI, J. 2014. Endoplasmic reticulum-mediated protein quality control in Arabidopsis. *Front Plant Sci,* 5**,** 162.

LIU, K., YUAN, C., FENG, S., ZHONG, S., LI, H., ZHONG, J. *et al.*, 2017a. Genome-wide analysis and characterization of Aux/IAA family genes related to fruit ripening in papaya (*Carica papaya L.*). BMC Genomics, 18, 351.

LIU, L., JIANG, Y., ZHANG, X., WANG, X., WANG, Y., HAN, Y. *et al.,* 2017b. Two SUMO proteases SUMO PROTEASE RELATED TO FERTILITY 1 and 2 are required for fertility in Arabidopsis. *Plant Physiology*, 175, 1703-1719.

LIU, Z. N., MIAO, L. M., HUO, R. X., SONG, X. Y., JOHNSON, C., KONG, L. J., SUNDARESAN, V. & YU, X. L. 2018. ARF2-ARF4 and ARF5 are Essential for Female and Male Gametophyte Development in Arabidopsis. *Plant and Cell Physiology,* 59**,** 179-189.

MACKAY, T. F. 2001. The genetic architecture of quantitative traits. *Annu Rev Genet,* 35**,** 303-39.

MCCALLUM, S., CULLEN, D. W., HEDLEY, P., SIMPSON, C., HANCOCK, R. D., MILNE, L., HACKETT, C. A. & GRAHAM, J. 2013. Control of fruit quality traits in red raspberry, a review from James Hutton Institute. *Fruits & roots: A celebration and forward look.* East Malling Research, East Malling, Kent, UK: Association of Applied Biologists.

MCCALLUM, S., SIMPSON, C. & GRAHAM, J. 2018. QTL Mapping and Marker Assisted Breeding in Rubus spp. London.

MCCALLUM, S., WOODHEAD, M., HACKETT, C. A., KASSIM, A., PATERSON, A. & GRAHAM, J. 2010. Genetic and environmental effects influencing fruit colour and QTL analysis in raspberry. *Theoretical and Applied Genetics,* 121**,** 611-627.

MCGINNIS, K. M., THOMAS, S. G., SOULE, J. D., STRADER, L. C., ZALE, J. M., SUN, T. P. & STEBER, C. M. 2003. The Arabidopsis SLEEPY1 gene encodes a putative F-box subunit of an SCF E3 ubiquitin ligase. *Plant Cell,* 15**,** 1120-1130.

MEESTERS, P. & PITSIOUDIS, A. 1999. Extention of the harvest period of raspberries in Belgium (moderate maritime climate). *Proceedings of the Seventh International Robus-Ribes Symposium***,** 419-421.

MENG, R. G. & FINN, C. 2002. Determining ploidy level and nuclear DNA content in Rubus by flow cytometry. *Journal of the American Society for Horticultural Science,* 127**,** 767-775.

MO, P., ZHU, Y., LIU, X., ZHANG, A., YAN, C. & WANG, D. 2007. Identification of two phosphatidylinositol/phosphatidylcholine transfer protein genes that are predominately transcribed in the flowers of Arabidopsis thaliana. *Journal of Plant Physiology,* 164, 478-486.

MOORE, P. P., PERKINS-VEAZIE, P., WEBER, C. A. & HOWARD, L. 2008. Environmental effect on antioxidant content of ten raspberry cultivars. *Proceedings of the Ixth International Rubus and Ribes Symposium*, 499-+.

MOOSE, S. P. & MUMM, R. H. 2008. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiology,* 147, 969-977.

MORGANTE, M. & SALAMINI, F. 2003. From plant genomics to breeding practice. *Curr Opin Biotechnol,* 14, 214-9.

MOYER, R., HUMMER, K., WROLSTAD, R. E. & FINN, C. 2002. Antioxidant compounds in diverse Ribes and Rubus germplasm. *Proceedings of the Eighth International Rubus and Ribes Symposium, Vols 1 and 2*, 501-505.

MUKHERJEE, D. & AHMAD, R. 2018. Resveratrol attenuates Nitrosodiethylamine-induced liver injury in anti-inflammatory manner via regulating cyclooxygenase-2. *Journal of Food Biochemistry,* 42.

MULLER, A., DUCHTING, P. & WEILER, E. W. 2002. A multiplex GC-MS/MS technique for the sensitive and quantitative single-run analysis of acidic phytohormones and related compounds, and its application to Arabidopsis thaliana. *Planta,* 216, 44-56.

NAGPAL, P., ELLIS, C. M., WEBER, H., PLOENSE, S. E., BARKAWI, L. S., GUILFOYLE, T. J., HAGEN, G., ALONSO, J. M., COHEN, J. D., FARMER, E. E., ECKER, J. R. & REED, J. W. 2005. Auxin response factors ARF6 and ARF8 promote jasmonic acid production and flower maturation. *Development,* 132, 4107-4118.

NEPI, M., GRASSO, D. A. & MANCUSO, S. 2018. Nectar in Plant-Insect Mutualistic Relationships: From Food Reward to Partner Manipulation. *Frontiers in Plant Science,* 9.

NIELSEN, A., REITAN, T., RINVOLL, A. W. & BRYSTING, A. K. 2017. Effects of competition and climate on a crop pollinator community. *Agriculture Ecosystems & Environment,* 246, 253-260.

OBROUCHEVA, N. V. 2014. Hormonal regulation during plant fruit development. *Russian Journal of Developmental Biology,* 45, 11-21.

OSTERGAARD, L., LAUVERGEAT, V. V., NAESTED, H., MATTSSON, O. & MUNDY, J. 2001. Two differentially regulated Arabidopsis genes define a new branch of the DFR superfamily. *Plant Sci,* 160, 463-472.

OZGA, J. A. & REINECKE, D. M. 2003. Hormonal interactions in fruit development. *Journal of Plant Growth Regulation,* 22**,** 73-81.

OZGA, J. A., YU, J. & REINECKE, D. M. 2003. Pollination-, development-, and auxin-specific regulation of gibberellin 3 beta-hydroxylase gene expression in pea fruit and seeds. *Plant Physiology,* 131**,** 1137-1146.

PAGNUSSAT, G. C., YU, H. J., NGO, Q. A., RAJANI, S., MAYALAGU, S., JOHNSON, C. S., CAPRON, A., XIE, L. F., YE, D. & SUNDARESAN, V. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in Arabidopsis. *Development,* 132**,** 603-14.

PAN, X. Q., WELTI, R. & WANG, X. M. 2008. Simultaneous quantification of major phytohormones and related compounds in crude plant extracts by liquid chromatography-electrospray tandem mass spectrometry. *Phytochemistry,* 69**,** 1773-1781.

PANDOLFINI, T., MOLESINI, B. & SPENA, A. 2007. Molecular dissection of the role of auxin in fruit initiation. *Trends in Plant Science,* 12**,** 327-329.

PATERSON, A. H. 1996b. Mapping genes responsible for differences in phenotype. *In:* PATERSON, A. H. (ed.) *Genome Mapping in Plants.* R.G. Landes Co. .

PATERSON, A., KASSIM, A., MCCALLUM, S., WOODHEAD, M., SMITH, K., ZAIT, D. & GRAHAM, J. 2013. Environmental and seasonal influences on red raspberry flavour volatiles and identification of quantitative trait loci (QTL) and candidate genes. *Theoretical and Applied Genetics,* 126**,** 33-48.

PATTISON, J. A., SAMUELIAN, S. K. & WEBER, C. A. 2007. Inheritance of Phytophthora root rot resistance in red raspberry determined by generation means and molecular linkage analysis. *Theoretical and Applied Genetics,* 115**,** 225-236.

PÉREZ-LLORCA, M., MUÑOZ, P., MÜLLER, M. and MUNNÉ-BOSCH, S. 2019. Biosynthesis, metabolism and function of auxin, salicylic acid and melatonin in climacteric and non-climacteric fruit. *Frontiers in Plant Science*, 10, 136.

PIESLINGER, A. M., HOEPFLINGER, M. C. & TENHAKEN, R. 2010. Cloning of Glucuronokinase from Arabidopsis thaliana, the last missing enzyme of the myo-inositol oxygenase pathway to nucleotide sugars. *J Biol Chem,* 285**,** 2902-10.

POMARES-VICIANA, T., DIE, J., DEL RIO-CELESTINO, M., ROMAN, B. & GOMEZ, P. 2017. Auxin signalling regulation during induced and parthenocarpic fruit set in zucchini. *Molecular Breeding,* 37.

POSE, D. & BOTELLA, M. A. 2009. Analysis of the arabidopsis dry2/sqe1-5 mutant suggests a role for sterols in signaling. *Plant Signal Behav,* 4**,** 873-4.

PRITTS, M. 2008. Primocane-fruiting raspberry production. *Hortscience,* 43**,** 1640-1641.

QIN, Y., LEYDON, A. R., MANZIELLO, A., PANDEY, R., MOUNT, D., DENIC, S., VASIC, B., JOHNSON, M. A. & PALANIVELU, R. 2009. Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil. *PLoS Genet,* 5**,** e1000621.

QUITO-AVILA, D. F., LIGHTLE, D. & MARTIN, R. R. 2014. Effect of Raspberry bushy dwarf virus, Raspberry leaf mottle virus, and Raspberry latent virus on Plant Growth and Fruit Crumbliness in 'Meeker' Red Raspberry. *Plant Disease,* 98**,** 176-183.

RAJON, E. & PLOTKIN, J. B. 2013. The evolution of genetic architectures underlying quantitative traits. *Proc Biol Sci,* 280**,** 20131552.

RAMIREZ, F. & DAVENPORT, T. L. 2013. Apple pollination: A review. *Scientia Horticulturae,* 162**,** 188-203.

ROBERTSON, K. R. 1974. Genera of Rosaceae in Southeastern United-States. *Journal of the Arnold Arboretum,* 55**,** 303-332.

ROHDICH, F., KIS, K., BACHER, A. & EISENREICH, W. 2001. The non-mevalonate pathway of isoprenoids: genes, enzymes and intermediates. *Current Opinion in Chemical Biology,* 5**,** 535-540.

ROTH, O., ALVAREZ, J. P., LEVY, M., BOWMAN, J. L., ORI, N. & SHANI, E. 2018. The KNOXI Transcription Factor SHOOT MERISTEMLESS Regulates Floral Fate in Arabidopsis. *Plant Cell,* 30**,** 1309-1321.

ROY, R., SCHMITT, A. J., THOMAS, J. B. & CARTER, C. J. 2017. Review: Nectar biology: From molecules to ecosystems. *Plant Sci,* 262**,** 148-164.

RUAN, Y. L., PATRICK, J. W., BOUZAYEN, M., OSORIO, S. & FERNIE, A. R. 2012. Molecular regulation of seed and fruit set. *Trends in Plant Science,* 17**,** 656-665.

SABLOWSKI, R. & DORNELAS, M. C. 2014. Interplay between cell growth and cell cycle in plants. *Journal of Experimental Botany,* 65**,** 2703-2714.

SAEZ, A., MORALES, C. L., RAMOS, L. Y. & AIZEN, M. A. 2014. Extremely frequent bee visits increase pollen deposition but reduce drupelet set in raspberry. *Journal of Applied Ecology,* 51**,** 1603-1612.

SANTNER, A., CALDERON-VILLALOBOS, L. I. A. & ESTELLE, M. 2009. Plant hormones are versatile chemical regulators of plant growth. *Nature Chemical Biology,* 5**,** 301-307.

SARGENT, D. J., RYS, A., NIER, S., SIMPSON, D. W. & TOBUTT, K. R. 2007. The development and mapping of functional markers in Fragaria and their transferability and potential for mapping in other genera. *Theoretical and Applied Genetics,* 114**,** 373-384.

SATO, R. & MAESHIMA, M. 2019. The ER-localized aquaporin SIP2;1 is involved in pollen germination and pollen tube elongation in Arabidopsis thaliana. *Plant Mol Biol,* 100**,** 335-349.

SCHROEDER, A., MUELLER, O., STOCKER, S., SALOWSKY, R., LEIBER, M., GASSMANN, M., LIGHTFOOT, S., MENZEL, W., GRANZOW, M. & RAGG, T. 2006. The RIN: an RNA integrity number for assigning integrity values to RNA measurements. BMC Mol Biol, 7, 3.

SCOFIELD, S., DEWITTE, W. & MURRAY, J. A. H. 2007. The KNOX gene SHOOT MERISTEMLESS is required for the development of reproductive meristematic tissues in Arabidopsis. *Plant Journal,* 50**,** 767-781.

SEBASTIAN, D., FERNANDO, F. D., RAUL, D. G. & GABRIELA, G. M. 2020. Overexpression of Arabidopsis aspartic protease APA1 gene confers drought tolerance. *Plant Sci,* 292**,** 110406.

SEIFERT, G. J. 2004. Nucleotide sugar interconversions and cell wall biosynthesis: how to bring the inside to the outside. *Curr Opin Plant Biol,* 7**,** 277-84.

SERRANI, J. C., RUIZ-RIVERO, O., FOS, M. & GARCIA-MARTINEZ, J. L. 2008. Auxin-induced fruit-set in tomato is mediated in part by gibberellins. *Plant Journal,* 56**,** 922-934.

SHRIVASTAVA, A. & GUPTA, V. B. 2011. Methods for the determination of limit of detection and limit of quantification of the analytical methods. *Chronicles of Young Scientists,* 2**,** 5.

SHU, K., WENGUAN, Z., FENG, C., XIAFENG, L. & WENYU, Y. 2018. Abscisic acid and gibberellins antagonistically mediate plant development and abiotic stress responses. *Frontiers in Plant Science*, 9, 416.

SIMON, S., RUHL, M., DE MONTAIGU, A., WOTZEL, S. & COUPLAND, G. 2015. Evolution of CONSTANS Regulation and Function after Gene Duplication Produced a Photoperiodic Flowering Switch in the Brassicaceae. *Molecular Biology and Evolution,* 32**,** 2284-2301.

SIMPSON, C. G., CULLEN, D. W., HACKETT, C. A., SMITH, K., HALLETT, P. D., MCNICOL, J., WOODHEAD, M. & GRAHAM, J. 2017. Mapping and expression of genes associated with raspberry fruit ripening and softening. *Theor Appl Genet,* 130**,** 557-572.

SIMURA, J., ANTONIADI, I., SIROKA, J., TARKOWSKA, D., STRNAD, M., LJUNG, K. & NOVAK, O. 2018. Plant Hormonomics: Multiple Phytohormone Profiling by Targeted Metabolomics. *Plant Physiol,* 177**,** 476-489.

SOCHOR, M., VASUT, R. J., SHARBEL, T. F. & TRAVNICEK, B. 2015. How just a few makes a lot: Speciation via reticulation and apomixis on example of European brambles (Rubus subgen. Rubus, Rosaceae). *Molecular Phylogenetics and Evolution,* 89**,** 13-27.

SONG, H. R., SONG, J. D., CHO, J. N., AMASINO, R. M., NOH, B. & NOH, Y. S. 2009. The RNA Binding Protein ELF9 Directly Reduces SUPPRESSOR OF OVEREXPRESSION OF CO1 Transcript Levels in Arabidopsis, Possibly via Nonsense-Mediated mRNA Decay. *Plant Cell,* 21**,** 1195-1211.

SONG, Y. H., LEE, I., LEE, S. Y., IMAIZUMI, T. & HONG, J. C. 2012. CONSTANS and ASYMMETRIC LEAVES 1 complex is involved in the induction of FLOWERING LOCUS T in photoperiodic flowering in Arabidopsis. *Plant Journal,* 69**,** 332-342.

STAFNE, E. T., CLARK, J. R., WEBER, C. A., GRAHAM, J. & LEWERS, K. S. 2005. Simple Sequence Repeat (SSR) Markers for Genetic Mapping of Raspberry and Blackberry. *Journal of the American Society for Horticultural Science,* 130**,** 722-728.

STEVE, L. 2011. *How to use statistics,* Gosport.

STINTZI, A., WEBER, H., REYMOND, P., BROWSE, J. & FARMER, E. E. 2001. Plant defense in the absence of jasmonic acid: the role of cyclopentenones. *Proc Natl Acad Sci U S A,* 98**,** 12837-42.

STRIK, B. C., CLARK, J. R., FINN, C. E. & BANADOS, M. P. 2007. Worldwide blackberry production. *Horttechnology,* 17**,** 205-213.

SUZUKI, M., NAKAGAWA, S., KAMIDE, Y., KOBAYASHI, K., OHYAMA, K., HASHINOKUCHI, H., KIUCHI, R., SAITO, K., MURANAKA, T. & NAGATA, N. 2009. Complete blockage of the mevalonate pathway results in male gametophyte lethality. *Journal of Experimental Botany,* 60**,** 2055-2064.

THOMPSON, M. M. 1997. Survey of chromosome numbers in Rubus (Rosaceae:Rosoideae). *Annals of the Missouri Botanical Garden,* 84**,** 128-164.

TIWARI, A., VIVIAN-SMITH, A., LJUNG, K., OFFRINGA, R. & HEUVELINK, E. 2013. Physiological and morphological changes during early and later stages of fruit growth in Capsicum annuum. *Physiologia Plantarum,* 147**,** 396-406.

TRAPP, M. A., DE SOUZA, G. D., RODRIGUES, E., BOLAND, W. & MITHOFER, A. 2014. Validated method for phytohormone quantification in plants. *Frontiers in Plant Science,* 5.

VAN OOIJEN, J. W. 2004. MapQTL ® 5 Software for the mapping of quantitative trait loci in experimental populations of diploid species. *Kyazma BV. Wageningen*.

VANNESTE, S. & FRIML, J. 2009. Auxin: a trigger for change in plant development. *Cell,* 136**,** 1005-16.

VANSTRAELEN, M. & BENKOVA, E. 2012. Hormonal Interactions in the Regulation of Plant Development. *Annual Review of Cell and Developmental Biology, Vol 28,* 28**,** 463-487.

VRIEZEN, W. H., FERON, R., MARETTO, F., KEIJMAN, J. & MARIANI, C. 2008. Changes in tomato ovary transcriptome demonstrate complex hormonal regulation of fruit set. *New Phytologist,* 177**,** 60-76.

VUJOVIC, T., RUZIC, D., CEROVIC, R., LEPOSAVIC, A., KARAKLAJIC-STAJIC, Z., MITROVIC, O. & ZURAWICZ, E. 2017. An assessment of the genetic integrity of micropropagated raspberry and blackberry plants. *Scientia Horticulturae,* 225**,** 454-461.

WADA, T., TACHIBANA, T., SHIMURA, Y. & OKADA, K. 1997. Epidermal cell differentiation in Arabidopsis determined by a Myb homolog, CPC. *Science,* 277**,** 1113-6.

WANG, Y., CHEN, Q., CHEN, T., TANG, H. R., LIU, L. & WANG, X. R. 2016. Phylogenetic Insights into Chinese Rubus (Rosaceae) from Multiple Chloroplast and Nuclear DNAs. *Frontiers in Plant Science,* 7.

WANG, Y., ZHANG, W. Z., SONG, L. F., ZOU, J. J., SU, Z. & WU, W. H. 2008. Transcriptome Analyses Show Changes in Gene Expression to Accompany Pollen Germination and Tube Growth in Arabidopsis. *Plant Physiology,* 148**,** 1201-1211.

WARD, J. A., BHANGOO, J., FERNANDEZ-FERNANDEZ, F., MOORE, P., SWANSON, J. D., VIOLA, R., VELASCO, R., BASSIL, N., WEBER, C. A. & SARGENT, D. J. 2013. Saturated linkage map construction in Rubus idaeus using genotyping by sequencing and genome-independent imputation. *Bmc Genomics,* 14.

WASTERNACK, C. & KOMBRINK, E. 2010. Jasmonates: Structural Requirements for Lipid-Derived Signals Active in Plant Stress Responses and Development. *Acs Chemical Biology,* 5**,** 63-77.

WATERS, D. L., HOLTON, T. A., ABLETT, E. M., LEE, L. S. & HENRY, R. J. 2005. cDNA microarray analysis of developing grape (Vitis vinifera cv. Shiraz) berry skin. *Funct Integr Genomics,* 5**,** 40-58.

WEBER, C. A., PERKINS-VEAZIE, P., MOORE, P. P. & HOWARD, L. 2008. Variability of antioxidant content in raspberry germplasm. *Proceedings of the Ixth International Rubus and Ribes Symposium***,** 493-+.

WEILER, E. W. 1984. Immunoassay of Plant-Growth Regulators. *Annual Review of Plant Physiology and Plant Molecular Biology,* 35**,** 85-95.

WEINTRAUB, P. G. & BEANLAND L. 2006. Insect vectors of phytoplasmas. *Annu Rev Entomol.* 51, 91-111. https://doi: 10.1146/annurev.ento.51.110104.151039.

WOODHEAD, M., MCCALLUM, S., SMITH, K., CARDLE, L., MAZZITELLI, L. & GRAHAM, J. 2008. Identification, characterisation and mapping of simple sequence repeat (SSR) markers from raspberry root and bud ESTs. *Molecular Breeding,* 22**,** 555-563.

WOODHEAD, M., WEIR, A., SMITH, K., MCCALLUM, S., MACKENZIE, K. & GRAHAM, J. 2010. Functional Markers for Red Raspberry. *Journal of the American Society for Horticultural Science,* 135**,** 418-427.

WOZNICKI, T. L., HEIDE, O. M., REMBERG, S. F. & SONSTEBY, A. 2016. Effects of controlled nutrient feeding and different temperatures during floral initiation on yield, berry size and drupelet numbers in red raspberry (Rubus idaeus L.). *Scientia Horticulturae,* 212**,** 148-154.

YAMAGUCHI, N., HUANG, J. B., XU, Y. F., TANOI, K. & ITO, T. 2017. Fine-tuning of auxin homeostasis governs the transition from floral stem cell maintenance to gynoecium formation. *Nature Communications,* 8.

YANG, X. H., XU, Z. H. & XUE, H. W. 2005. Arabidopsis membrane steroid binding protein 1 is involved in inhibition of cell elongation. *Plant Cell,* 17**,** 116-31.

YANG, Y., XU, R., MA, C. J., VLOT, A. C., KLESSIG, D. F. & PICHERSKY, E. 2008. Inactive methyl indole-3-acetic acid ester can be hydrolyzed and activated by several esterases belonging to the AtMES esterase family of Arabidopsis. *Plant Physiology,* 147**,** 1034-1045.

YU, J., PRESSOIR, G., BRIGGS, W. H., VROH BI, I., YAMASAKI, M., DOEBLEY, J. F., MCMULLEN, M. D., GAUT, B. S., NIELSEN, D. M., HOLLAND, J. B., KRESOVICH, S. & BUCKLER, E. S. 2006. A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet,* 38**,** 203-8.

ZAHARAH, S. S., SINGH, Z., SYMONS, G. M. & REID, J. B. 2012. Role of Brassinosteroids, Ethylene, Abscisic Acid, and Indole-3-Acetic Acid in Mango Fruit Ripening. *Journal of Plant Growth Regulation,* 31**,** 363-372.

ZAVESKA DRABKOVA, L. & HONYS, D. 2017. Evolutionary history of callose synthases in terrestrial plants with emphasis on proteins involved in male gametophyte development. *PLoS One,* 12**,** e0187331.

ZHANG, M., YUAN, B. & LENG, P. 2009. The role of ABA in triggering ethylene biosynthesis and ripening of tomato fruit. *Journal of Experimental Botany,* 60**,** 1579-1588.

ZWANENBURG, B., POSPISIL, T. & ZELJKOVIC, S. C. 2016. Strigolactones: new plant hormones in action. *Planta,* 243**,** 1311-1326.

Appendix:

Crumbly' microarray probe	***A. thaliana* gene ID**	**G.O. annotations**
CUST_50855_PI426541283	AT1G25540.2	flower development
CUST_19403_PI426541283	AT1G69120.1	flower development; floral meristem determinacy
CUST_52339_PI426541283	AT1G79730.1	negative regulation of flower development
CUST_14645_PI426541283	AT1G80940.1	regulation of flower development
CUST_50855_PI426541283	AT1G25540.2	flower development
CUST_55276_PI426541283	AT3G13730.1	stamen development - Brassinosteroid biosynthetic process - petal development
CUST_32226_PI426541283	AT3G12280.1	embryo sac development - double fertilization forming a zygote and endosperm - gametophyte development - endosperm development
CUST_34459_PI426541283	AT3G61120.1	plant ovule development
CUST_5790_PI426541283	AT5G05660.1	flower development
CUST_51962_PI426541283	AT5G14530.1	flower development
CUST_34701_PI426541283	AT5G16260.1	negative regulation of flower development
CUST_27902_PI426541283	AT1G65650.1	shoot meristem development

CONTINUE NEXT PAGE

CUST_10154_PI426541283	AT3G63440.1	cytokinin metabolic process; cytokinin dehydrogenase activity
CUST_10338_PI426541283	AT3G48170.1	response to Abscisic Acid (ABA)
CUST_12078_PI426541283	AT1G17840.1	response to ABA
CUST_12755_PI426541283	AT1G76490.1	sterol biosynthetic process
CUST_13398_PI426541283	AT2G04240.1	ABA metabolic process; response to Gibberellins
CUST_1495_PI426541283	AT4G21410.1	response to ABA
CUST_15482_PI426541283	AT1G73805.1	regulation of SA biosynthetic process
CUST_1641_PI426541283	AT3G48690.1	methyl indole-3- acetate (MeIAA) acetate esterase activity
CUST_16869_PI426541283	AT2G17840.1	response to ABA
CUST_17274_PI426541283	AT5G07700.1	response to jasmonic acid (JA)/Auxin/salicylic acid (SA)
CUST_18003_PI426541283	AT4G24520.1	response to ABA
CUST_18787_PI426541283	AT2G04160.1	response to Auxin
CUST_18959_PI426541283	AT4G24210.1	Gibberellins mediated signalling pathway
CUST_20475_PI426541283	AT1G55020.1	response to JA, ABA; JA biosynthetic process
CUST_20489_PI426541283	AT1G11910.1	response cytokinin

CONTINUE NEXT PAGE

CUST_22099_PI426541283	AT5G52240.1	steroid binding
CUST_2356_PI426541283	AT5G40990.1	response to SA
CUST_24740_PI426541283	AT4G22540.1	sterol binding; signal transduction; sterol transporter activity
CUST_26909_PI426541283	AT2G43820.1	SA glucosyl transpherase activity; SA metabolic process
CUST_2783_PI426541283	AT2G22570.2	response to ABA
CUST_27978_PI426541283	AT1G15780.1	response to SA
CUST_28007_PI426541283	AT3G29770.1	MeIAA, MeJA and MeSA esterase activity: SA and JA metabolic process
CUST_29488_PI426541283	AT2G34680.1	response to auxin
CUST_33454_PI426541283	AT1G05180.1	Auxin activated signalling pathway; auxin homeostasis; response to cytokinin
CUST_3392_PI426541283	AT3G53260.1	Cinnamic Acid (CA) biosynthetic process
CUST_35443_PI426541283	AT1G58440.1	sterol biosynthetic process
CUST_36145_PI426541283	AT2G01530.1	response to cytokinin
CUST_37473_PI426541283	AT5G51300.3	response to ABA
CUST_37776_PI426541283	AT5G25610.1	response to ABA

CONTINUE NEXT PAGE

CUST_37835_PI426541283	AT1G30330.2	response to Auxin; auxin-activated signalling pathway
CUST_38364_PI426541283	AT1G53130.1	regulation of JA and SA biosynthetic process
CUST_38657_PI426541283	AT2G46410.1	response to JA and SA
CUST_39993_PI426541283	AT1G77120.1	response to ABA
CUST_40382_PI426541283	AT3G51550.1	response to ethylene - response to brassinosteroids - ABA and Br activated signalling pathway
CUST_40511_PI426541283	AT1G48910.1	indole-3-pyruvate (IPyA) is the major IAA biosynthesis intermediate
CUST_41772_PI426541283	AT3G03050.1	response to OPDA
CUST_42274_PI426541283	AT1G19190.1	MeIAA esterase activity
CUST_42319_PI426541283	AT3G02590.1	sterol biosynthetic process
CUST_42548_PI426541283	AT2G33590.1	response to ABA
CUST_44536_PI426541283	AT2G03440.1	ABA activated signalling pathway
CUST_44619_PI426541283	AT3G12110.1	response to OPDA
CUST_44664_PI426541283	AT2G33150.1	JA biosynthetic processes
CUST_46340_PI426541283	AT2G16385.1	hormone activity

CONTINUE NEXT PAGE

CUST_4664_PI426541283	AT2G03550.1	MeIAA- acetate esterase activity
CUST_46817_PI426541283	AT3G02875.1	IAA-Phe-Leu conjugate hydrolase activity
CUST_47833_PI426541283	AT5G36140.1	brassinosteroid homeostasis/brassinisteroids biosynthesis process
CUST_48849_PI426541283	AT5G49570.1	response to SA
CUST_34797_PI426541283	AT1G14830.1	pollen maturation
CUST_1684_PI426541283	AT1G32310.1	pollen development
CUST_14357_PI426541283	AT1G34300.1	recognition of pollen
CUST_19171_PI426541283	AT1G63180.1	pollen development
CUST_48564_PI426541283	AT1G63530.1	pollen tube growth
CUST_55235_PI426541283	AT1G77980.1	regulation of pollen tube growth
CUST_3771_PI426541283	AT2G13680.1	regulation of pollen tube growth; pollen germination; microsporogenesis; pollen wall assembly
CUST_51278_PI426541283	AT2G15890.1	pollen tube guidance
CUST_14032_PI426541283	AT2G31220.1	pollen development; anther wall tapetum development
CUST_41580_PI426541283	AT2G36880.2	pollen tube growth

CONTINUE NEXT PAGE

CUST_43359_PI426541283	AT3G01640.1	pollen tube development
CUST_29550_PI426541283	AT3G13220.1	pollen exine formation
CUST_55276_PI426541283	AT3G13730.1	stamen development - Brassinosteroid biosynthetic process
CUST_40382_PI426541283	AT3G51550.1	pollen tube reception
CUST_18020_PI426541283	AT4G04970.1	pollen development
CUST_7407_PI426541283	AT4G24973.1	rejection of self pollen
CUST_26212_PI426541283	AT4G30520.1	anther development
CUST_40952_PI426541283	AT5G15650.1	pollen development
CUST_33524_PI426541283	AT5G20690.1	pollen tube guidance
CUST_20976_PI426541283	AT5G48230.2	pollen tube growth
CUST_10849_PI426541283	AT2G41190.1	amino acid transmembrane transport
CUST_26851_PI426541283	AT5G53130.1	calcium iona transport
CUST_16398_PI426541283	AT2G04620.1	cation transmembrane transporter activity
CUST_2684_PI426541283	AT5G24390.1	intracellular protein transport

CONTINUE NEXT PAGE

CUST_33876_PI426541283	AT5G48485.1	lipid transporter activity
CUST_35273_PI426541283	AT1G11260.1	monosaccharide transmembrane transport - glucose import
CUST_13929_PI426541283	AT1G79400.1	monovalent cation:proton antiporter activity
CUST_36009_PI426541283	AT1G69870.1	peptide transmembrane transporter activity - nitrate transport
CUST_2161_PI426541283	AT2G18840.1	polysaccharide transport
CUST_33654_PI426541283	AT3G02050.1	potassium ion transmembrane transporter activity
CUST_34391_PI426541283	AT3G51570.1	signal transduction
CUST_45820_PI426541283	AT5G22400.1	signal transduction
CUST_40022_PI426541283	AT1G51610.1	transmembrane transport - efflux transmembrane transporter activity
CUST_12940_PI426541283	AT2G36380.1	transmembrane transport - ATPase-coupled transmembrane transporter activity
CUST_42668_PI426541283	AT1G03900.1	transmembrane transport - ATPase-coupled transmembrane transporter activity
CUST_3482_PI426541283	AT1G76670.1	transmembrane transport
CUST_15656_PI426541283	AT3G21690.1	transmembrane transport
CUST_50866_PI426541283	AT3G21250.1	transmembrane transport
CUST_43040_PI426541283	AT5G17630.1	transmembrane transporter activity
CUST_32113_PI426541283	AT3G56950.1	transmembrane transport

Table A.3.1: List of 107 'crumbly' microarray probes significantly differently expressed with respect to stage*phenotype interaction and matching R. idaeus genes whose A. thaliana ortholog have gene ontology annotation related to: flower development, hormone, pollen and transport.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID
CUST_33524_PI426541283	pollen tube guidance	GO:0010183	AT5G20690.1
CUST_32113_PI426541283	pollen germination	GO:0009846	AT3G56950.1
	pollen tube growth	GO:0009860	
	transmembrane transport	GO:0055085	
CUST_1684_PI426541283	pollen development	GO:0009555	AT1G32310.1
CUST_18020_PI426541283	pollen development	GO:0009555	AT4G04970.1
CUST_20976_PI426541283	pollen tube growth	GO:0009860	AT5G48230.1
CUST_55235_PI426541283	regulation of pollen tube growth	GO:0080092	AT1G77980.1
CUST_29550_PI426541283	pollen exine formation	GO:0010584	AT3G13220.1
CUST_34797_PI426541283	pollen maturation	GO:0010152	AT1G14830.1

Table A.3.2: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to pollen and not included into the two clusters of probes highlighted in tree cluster heatmap of Figure 3.3.

List of the eight microarray probes, significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significant at 95% confidence levels. The eight probes did not belong to any of the two clusters highlighted in the heatmap of Figure 3.6. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to pollen.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID
CUST_2356_PI426541283	response to salicylic acid (SA)	GO:0009751	AT4G24520.1
CUST_8434_PI426541283	response to jasmonic acid (JA)	GO:0009753	AT2G33590.1
CUST_36145_PI426541283	response to cytokinin (CK)	GO:0009735	AT2G17840.1
CUST_2783_PI426541283	response to abscisic acid (ABA)	GO:0009737	AT5G51300.3
CUST_55698_PI426541283	response to Auxin (AUX)	GO:0009733	AT1G77120.1
CUST_29488_PI426541283	response to Auxin (AUX)	GO:0009733	AT2G46410.1
	response to ethylene	GO:0009723	
CUST_40382_PI426541283	response to brassinosteroid (Br)	GO:0009741	AT1G15780.1

Table A.3.3: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to 'response to hormone' and not included into the two clusters of probes highlighted in tree cluster heatmap of Figure 3.6.

List of the seven microarray probes, significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significant at 95% confidence levels. The probes did not belong to any of the three clusters highlighted in the heatmap of Figure 3.6. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to 'response to hormone'.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID
CUST_40511_PI426541283	indole-3-pyruvate monooxygenase activity	GO:0103075	AT1G48910.1
	auxin biosynthetic process	GO:0009851	
CUST_38364_PI426541283	regulation of salicylic acid biosynthetic process	GO:0080142	AT1G53130.1
	regulation of jasmonic acid biosynthetic process	GO:0080141	
CUST_52236_PI426541283	methyl indole-3-acetate esterase activity	GO:0080030	AT5G14930.2
CUST_15482_PI426541283	regulation of salicylic acid biosynthetic process	GO:0080142	AT1G73805.1

Table A.3.4: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to 'hormone biosynthesis' and not included into the two clusters of probes highlighted in tree cluster heatmap of Figure 3.7.

List of the four microarray probes, significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significant at 95% confidence levels. The probes did not belong to any of the two clusters highlighted in the heatmap of Figure 3.7. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to 'hormone biosynthesis'.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID
CUST_49535_PI426541283	auxin biosynthetic process	GO:0010337	AT1G28380.1
CUST_46817_PI426541283	[a]IAA-Leu conjugate hydrolase activity	GO:0010211	AT3G02875.1
	[b]IAA-Phe conjugate hydrolase activity	GO:0010210	
	auxin metabolic process	GO:0009850	
CUST_26909_PI426541283	salicylic acid metabolic process	GO:0009696	AT2G43820.1
	benzoate metabolic process	GO:0018874	
	salicylic acid glucosyl transferase activity	GO:0052640	
	benzoic acid glucosyl transferase activity	GO:0052641	
	salicylic acid glucosyl transferase (ester forming) activity	GO:0052639	

[a]IAA-Leu – Indole-3-acetic acid-Luecine conjugated hydrolase activity
[b]IAA-Phe – Indole-3-acetic acid-Phenylalanine conjugated hydrolase activity

Table A.3.5: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to 'hormone other' and not included into the two clusters of probes highlighted in tree cluster heatmap of Figure 3.8.

List of the three microarray probes, significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significant at 95% confidence levels. The probes did not belong to any of the two clusters highlighted in the heatmap of Figure 3.8. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to the miscellaneous group 'hormone other'.

Microarray probe	Gene Ontology (GO) term	GO code	*A. thaliana* gene ID
CUST_50866_PI426541283	transmembrane transport	GO:0055085	AT1G28380.1
CUST_12940_PI426541283	transmembrane transport	GO:0055085	AT2G36380.1
	ATPase-coupled transmembrane transporter activity	GO:0042626	
CUST_45820_PI426541283	signal transduction	GO:0007165	AT5G22400.1
CUST_43040_PI426541283	transmembrane transporter activity	GO:0022857	AT5G17630.1
CUST_34391_PI426541283	signal transduction	GO:0007165	AT3G51570.1
CUST_35273_PI426541283	monosaccharide transmembrane transport	GO:0015749	AT1G11260.1
	glucose import	GO:0046323	
CUST_3482_PI426541283	transmembrane transport	GO:0055085	AT1G76670.1
CUST_33654_PI426541283	potassium ion transmembrane transporter activity	GO:0015079	AT3G02050.1
CUST_42668_PI426541283	transmembrane transport	GO:0055085	AT1G03900.1
	ATPase-coupled transmembrane transporter activity	GO:0042626	
CUST_15656_PI426541283	transmembrane transport	GO:0055085	AT3G21690.1
CUST_33876_PI426541283	lipid transporter activity	GO:0005319	AT5G48485.1

Table A.3.6: 'Crumbly' microarray probes matching *A. thaliana* genes with ontology terms related to 'transport' and not included into the two clusters of probes highlighted in tree cluster heatmap of Figure 3.9.

List of the eleven microarray probes, significantly differently expressed between the two phenotypes (i.e. mostly and never 'crumbly') with differences being significant at 95% confidence levels. The probes did not belong to any of the two clusters highlighted in the heatmap of Figure 3.9. The gene ID of the *A. thaliana* gene corresponding to the *R. idaeus* gene matched by the probe was reported together with the ontology term and its code; obviously related to 'transport'.

Pedigree					
♀ mother	**x**	**♂ father**	**genotype**		**phenotype**
DJ1185	**x**	8510A21	0560E11	FC	crumbly
Tulameen	**x**	8510A28	0565F3	FC	crumbly
Malling Minerva	**x**	8510A41	0663RE3	FC	crumbly
R4A1	**x**	Glen Fyne	0946/4	FC	crumbly
7826C1	**x**	8627RE7	9059D-2	FC	crumbly
9421A4	**x**	9434B-1	99105RC-2	FC	crumbly
9426C-5	**x**	9429E-2	99116E-4	FC	crumbly
SCRI 8631D-1	**x**	SCRI 8605C-2	Glen Fyne	FC	crumbly
9059D-2	**x**	8510A73	0019E2	FC	crumbly
9455F-2	**x**	8510A5	0304F6	FC	crumbly
0096RF-4	**x**	8510A6	04101A5	FC	crumbly
0003RB1	**x**	8510A9	0433F2	FC	crumbly
9046RA2	**x**	8510A22	0511F1	FC	crumbly
Glen Fyne	**x**	8510A29	0573B5	FC	crumbly
0312E3	**x**	8510A35	0658C5	FC	crumbly
0312E3	**x**	8510A32	0658E-1	FC	crumbly
0312E3	**x**	8510A37	0658F-7	FC	crumbly
Glen Ample	**x**	8510A43	0671D-4	FC	crumbly
97134B1	**x**	8510A57	0867E-4	FC	crumbly
R4A1	**x**	Glen Fyne	0946/19	FC	crumbly
8735J-7	**x**	8626RJ-2	9050RD3	FC	crumbly
8820E3	**x**	88K-7	9238D5	FC	crumbly
EM5961/1	**x**	7826C1	9350F3	FC	crumbly
WSU1068	**x**	ORUS 2078	97134B1	FC	crumbly
9349F5	**x**	9349A4	9764F-3	FC	crumbly
9351D-3	**x**	9350E1	9769RD1	FC	crumbly
	x		Autumn Bliss	PF	crumbly
Autumn Bliss[1]	**x**		Erika	PF	crumbly
7326E1	**x**	7412H6	Glen Ample	FC	crumbly
00123A5	**x**	0019B11	Glen Dee	FC	crumbly
SCRI 7331/1	**x**	SCRI 7256/1	Glen Lyon	FC	crumbly
SCRI 688/12	**x**	SCRI 6815/113	Glen Moy	FC	crumbly
SCRI 6531/84	**x**	SCRI 6549/1	Glen Prosen	FC	crumbly
7326E1	**x**	7412H16	Glen Rosa	FC	crumbly

[1]open pollinated - FC (floricane) – PF (primocaine)

CONTINUE NEXT PAGE

Pedigree					
♀ mother	**x**	**♂ father**	**genotype**		**phenotype**
7741D4	**x**	7919B11	Glen Shee	FC	crumbly
	x		Imara	PF	crumbly
Autumn Bliss	**x**		Joan Squire	PF	crumbly
	x		Kweli	PF	crumbly
King	**x**	Louden	Latham	FC	crumbly
Willamette	**x**	Cuthbert	Meeker	FC	crumbly
(unknown)	**x**		Obbard	PF	crumbly
Nootka	**x**	Glen Prosen	Tulameen	FC	crumbly
9431G-8	**x**	8510A71	Sanibelle	FC	no crumbly
0015D3	**x**	8510A14	0453C4	FC	no crumbly
	x		00123A7	FC	no crumbly
Glen Rosa	**x**	8510A20	0550E4	FC	no crumbly
0304F6	**x**	Autumn Treasure	0925B4	PF	no crumbly
0304F6	**x**	Autumn Treasure	0925B8	PF	no crumbly
0304F6	**x**	Autumn Treasure	0925D-15	PF	no crumbly
8003G10	**x**	8003C1	8605C-2	FC	no crumbly
8020E8	**x**	8631D-1	9025A1	FC	no crumbly
complex	**x**		Autumn Britten	PF	no crumbly
EM6304/36	**x**	EM6330/96	Autumn Treasure	PF	no crumbly
SCRI 9422C-4	**x**	SCRI 9434B-1	Glen Cally	FC	no crumbly
0030E-12	**x**	8510A16	Glen Carron	FC	no crumbly
Glen Rosa	**x**	SCRI 8605C-2	Glen Doll	FC	no crumbly
SCRI 9422C-4	**x**	SCRI 9434B-1	Glen Ericht	FC	no crumbly
	x		Malling Leo	FC	no crumbly
Joan Squire	**x**	complex	Brice	PF	crumbly
	x		Chief	FC	no crumbly
Joan J	**x**	complex	Joan Irene	PF	crumbly
EM3689	**x**	Gaia	Malling Hiesta	FC	crumbly
EM selection	**x**	SCRI selection	Malling Minerva	FC	no crumbly
Lloyd George	**x**	Preussen	Schoenemann	FC	crumbly
FC (floricane) – PF (primocaine)					

Table A.4.1: Full list of 63 selected genotypes for the validation pool of the 'crumbly' markers.

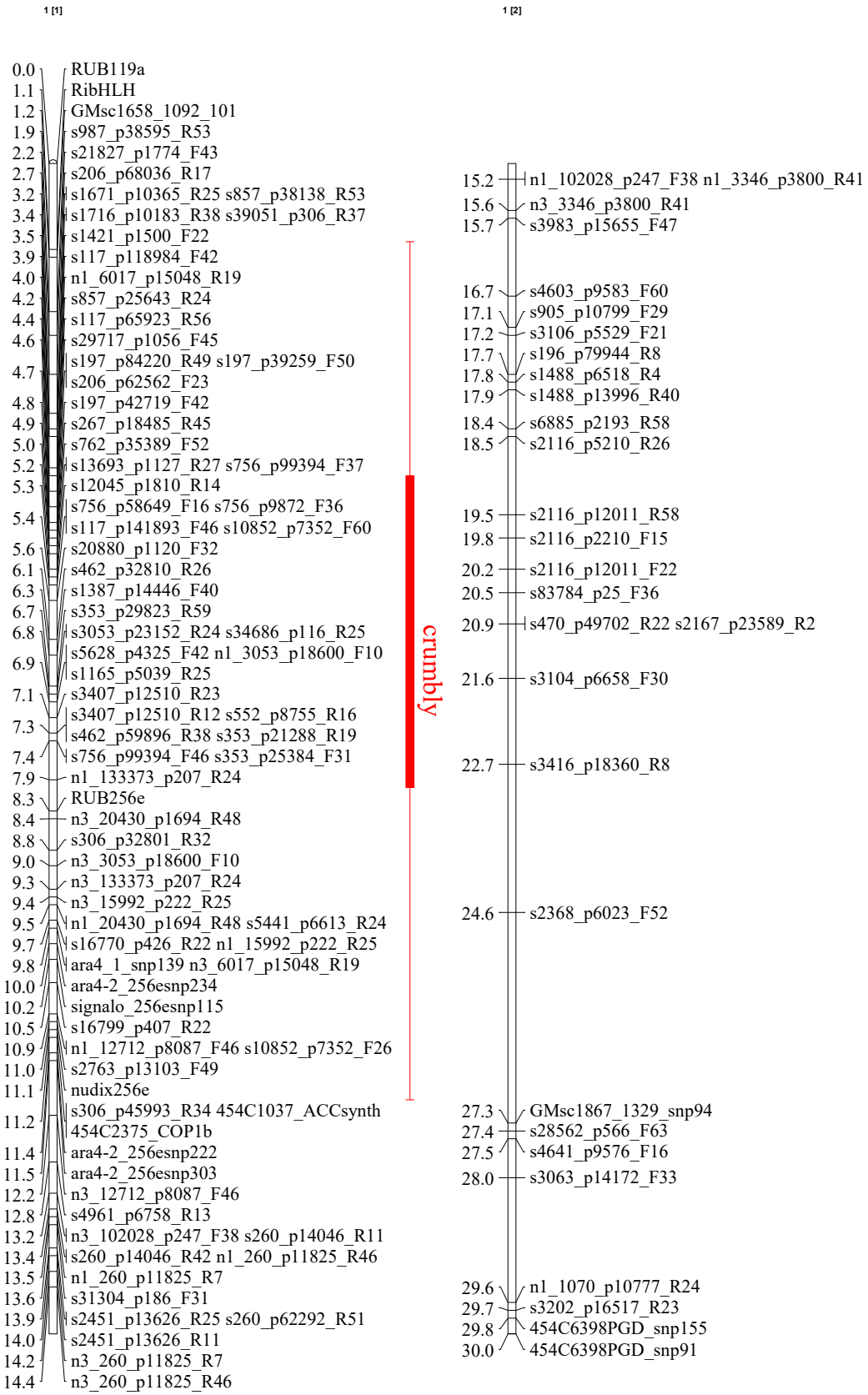
List of 63 different genotypes of which 45 specifically selected for their tendency to show 'crumbly' symptoms and 18 on the contrary for having never displayed the symptoms.

***A. thaliana* gene ID**	**Crumbly**	**stage expression levels**			**[a]S.E. crumbly/stage**	**Probe microarray**	**Linkage group**
		closed	**open**	**green**			
AT4G26840.1	mostly	0.038	-0.100	-0.165*	0.0395	CUST_6848_PI426541283	3[b]
	never	0.183	0.143	-0.084*			
AT5G08080.3	mostly	-0.0252*	-0.0415	-0.0100	0.0309	CUST_2371_PI426541283	3[b]
	never	0.0372*	0.1403	0.0138			
AT2G21540.1	mostly	0.241*	-0.103	-0.245*	0.0649	CUST_26373_PI426541283	3[c]
	never	0.159*	0.218	-0.302*			

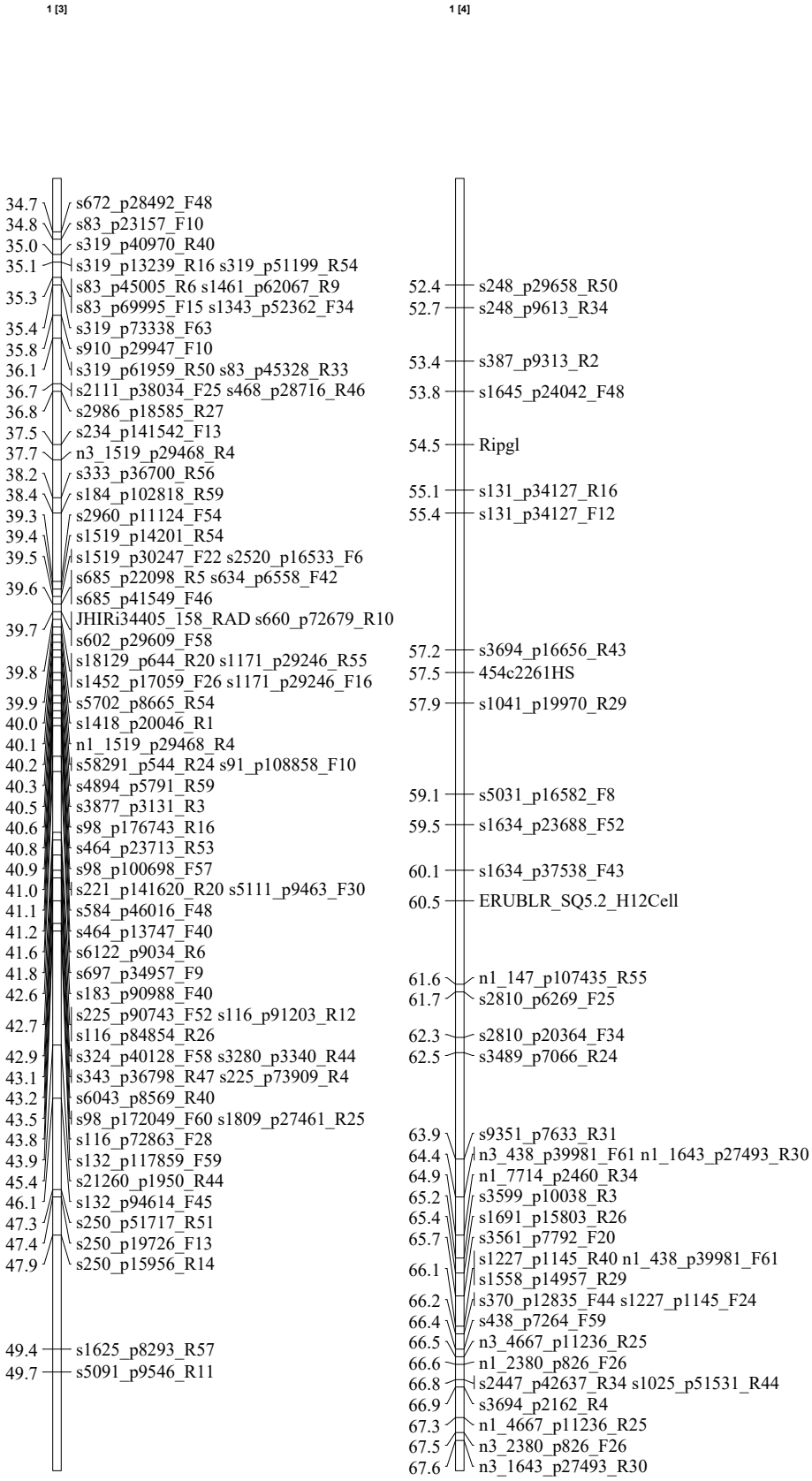
[a]2 degrees of freedom and 4 replicates
[b] 'crumbly' QTL on linkage group 3 previously identified by Graham et al., (2015)
[c] 'crumbly' QTL on linkage group 3 identified during this work
*differences not statistically significant at 99.9% confidence levels

Table A.6.1: ANOVA table of means (stage*phenotype interaction) of microarray probes mapped inside the two 'crumbly' QTLs on linkage group 3.

List of the probes matching *Arabidopsis thaliana* predicted genes with gene ontology (GO) annotation others than: flower development, hormones, pollen or transport. The probes mapped inside the two 'crumbly QTLs, one on linkage group 3. ANOVA table of means for the interaction phenotype (i.e. mostly and never 'crumbly') per stage (i.e. closed bud, open flower and green berry).



CONTINUE NEXT PAGE



CONTINUE NEXT PAGE

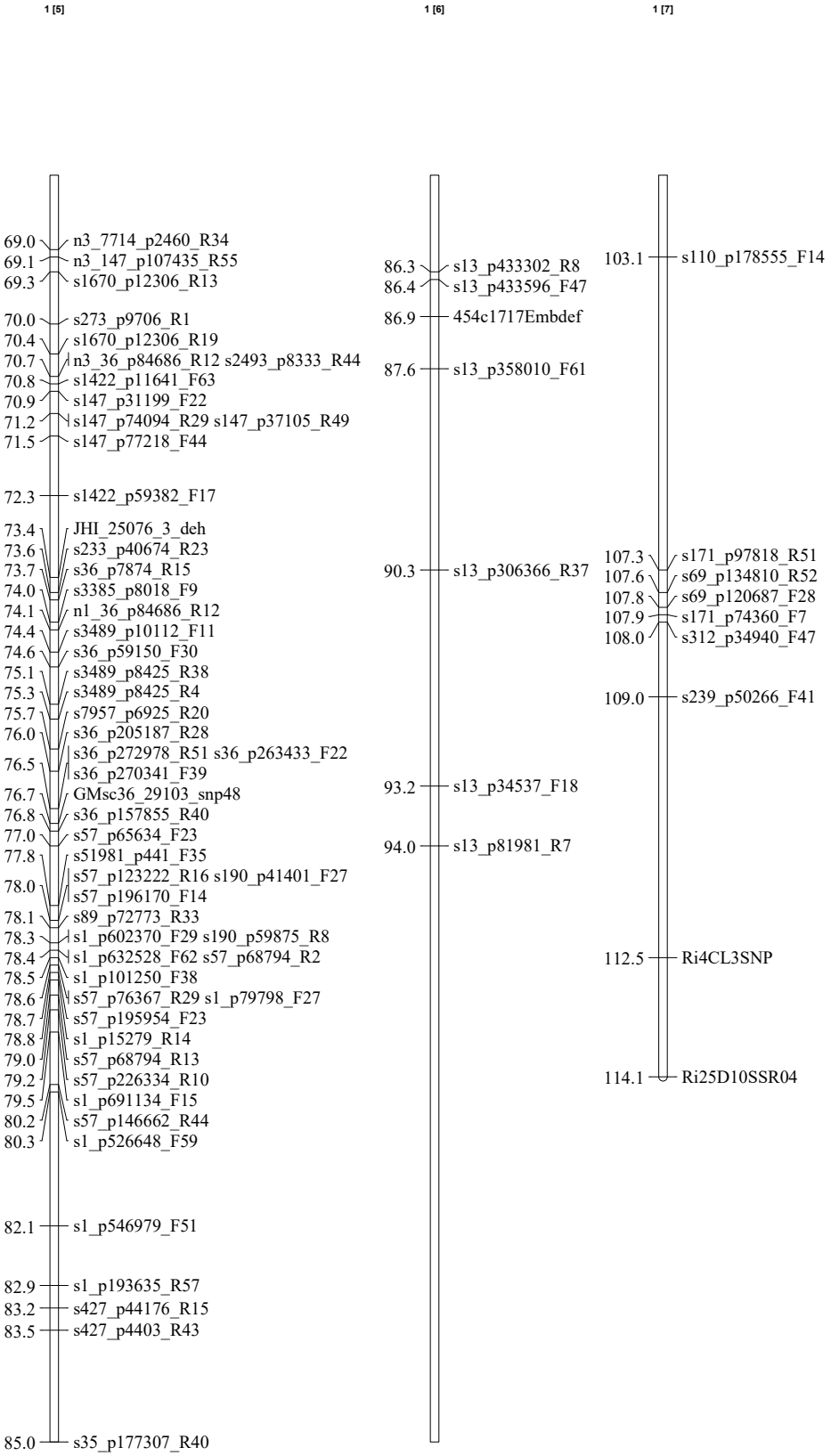
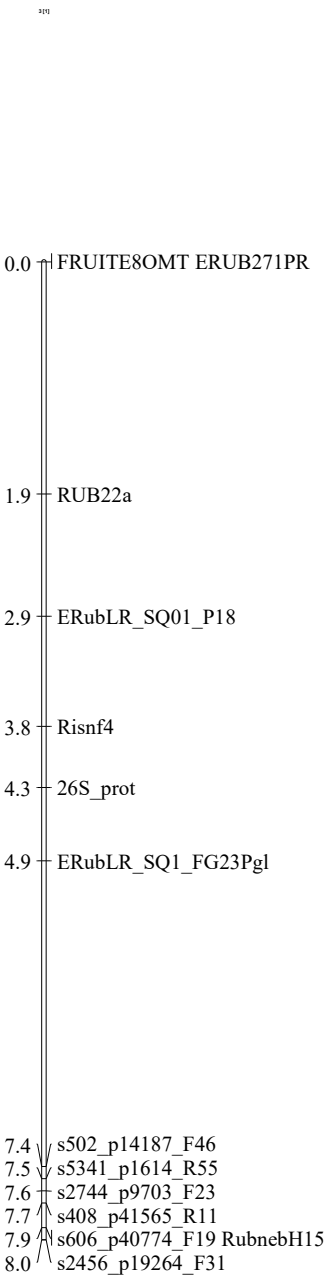
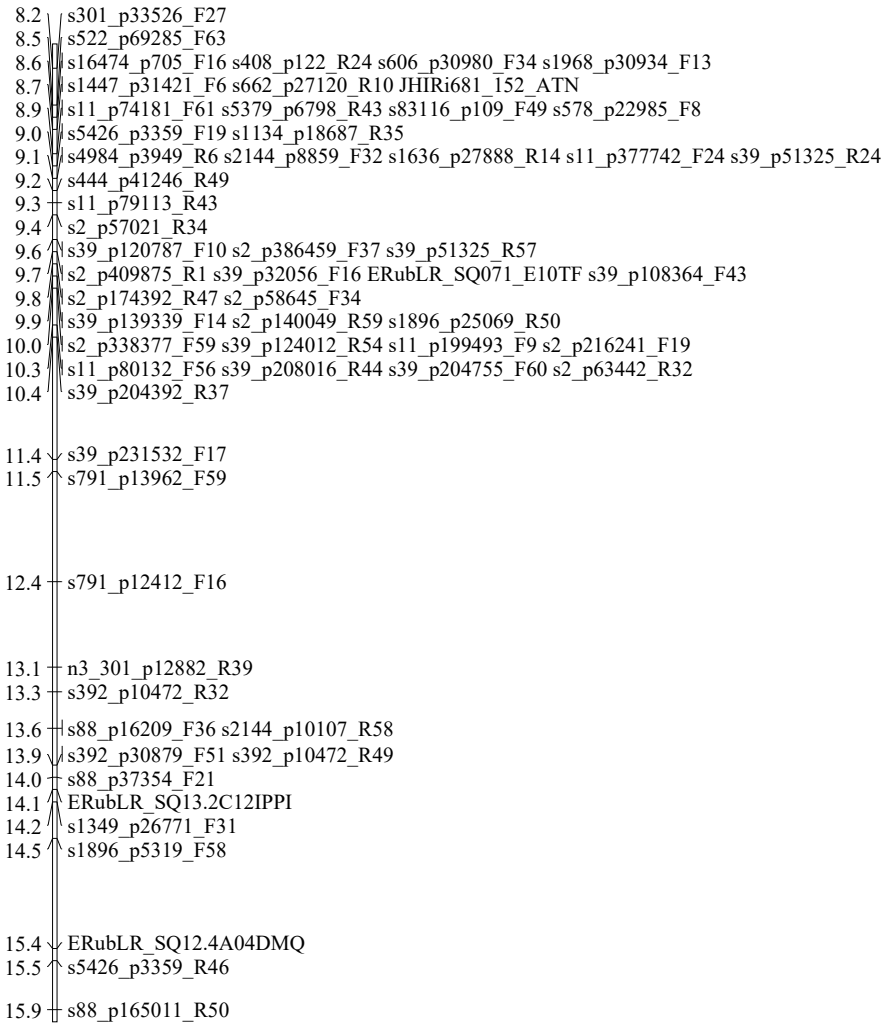


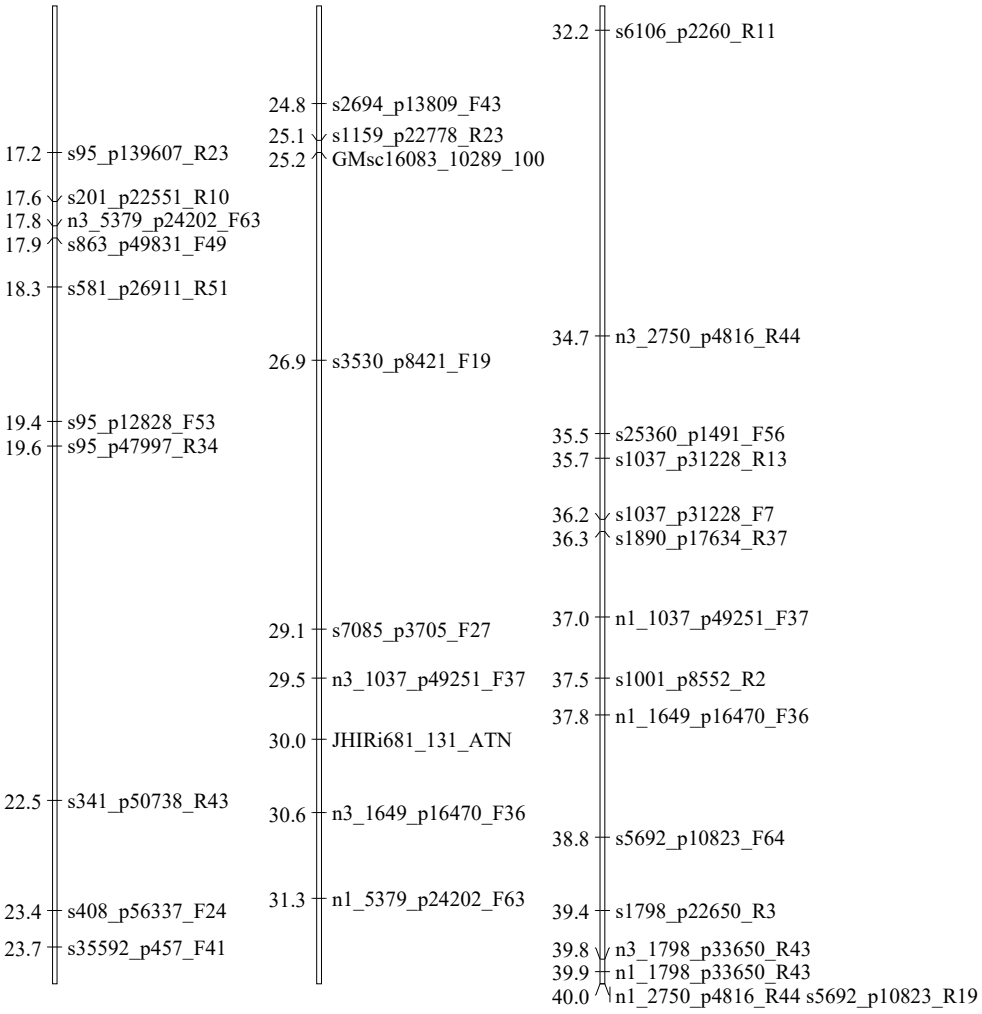
Figure A.4.1: **Lingage group 1 (LG1) full markers map and 'crumbly' QTL.**



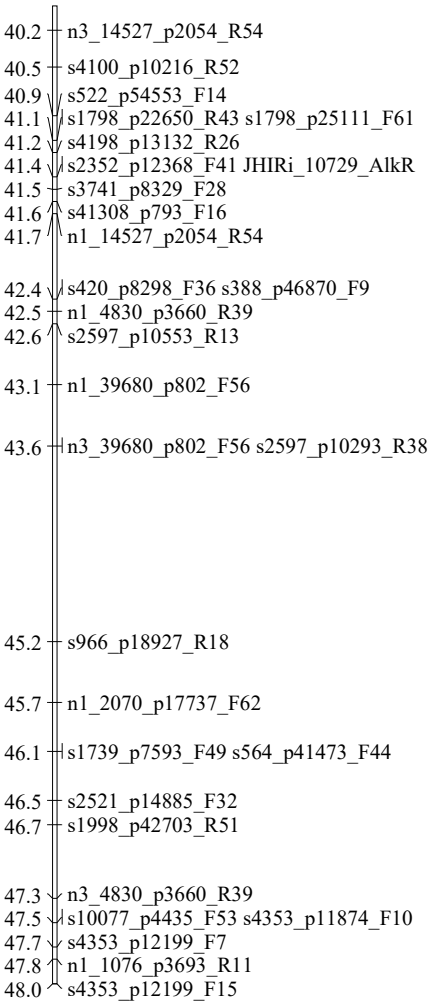
CONTINUE NEXT PAGE



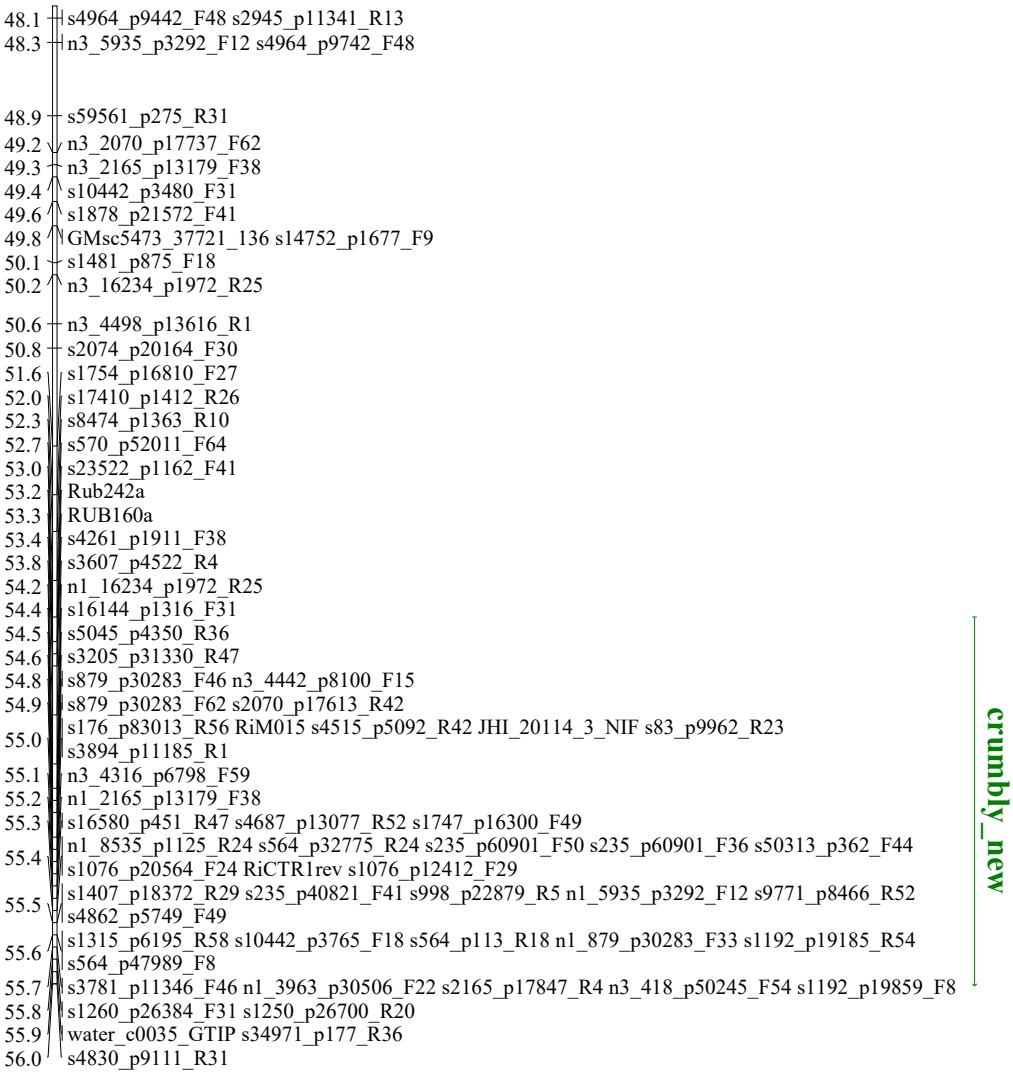
CONTINUE NEXT PAGE



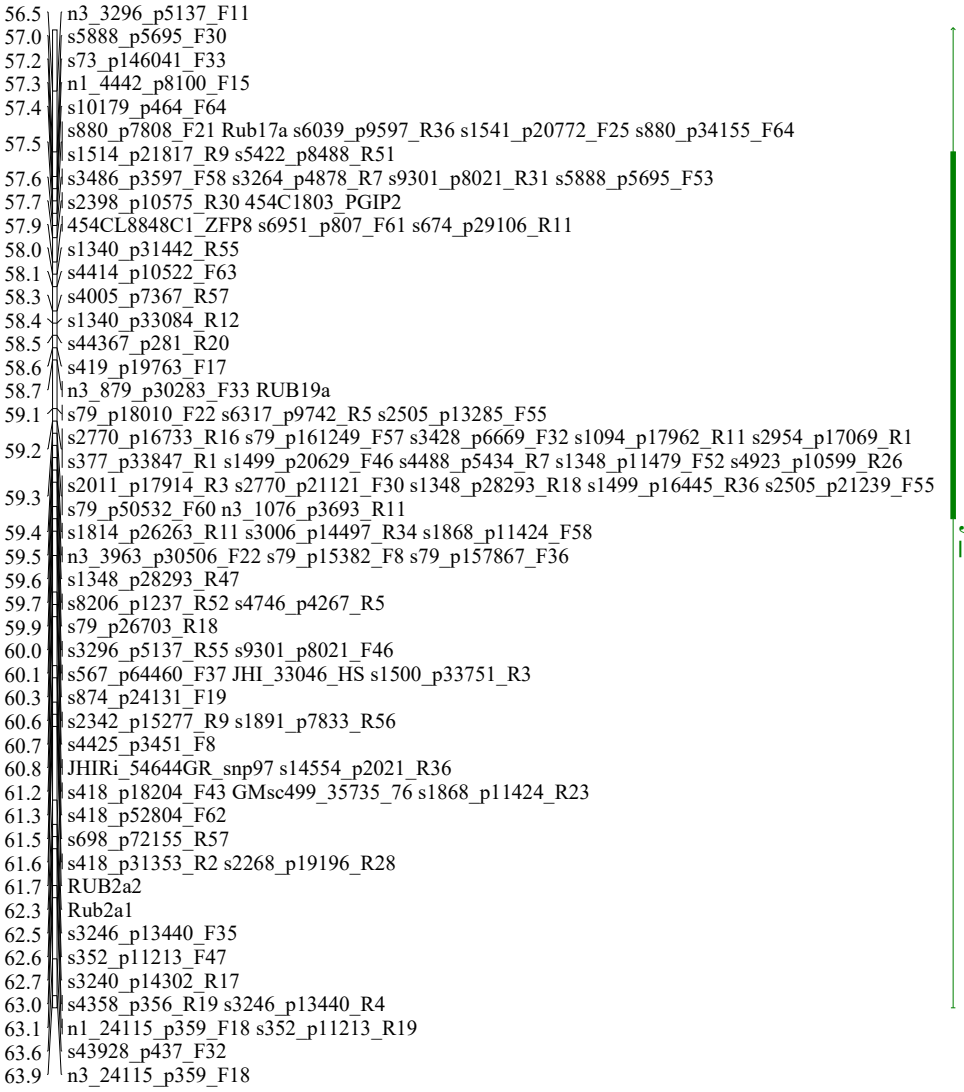
CONTINUE NEXT PAGE



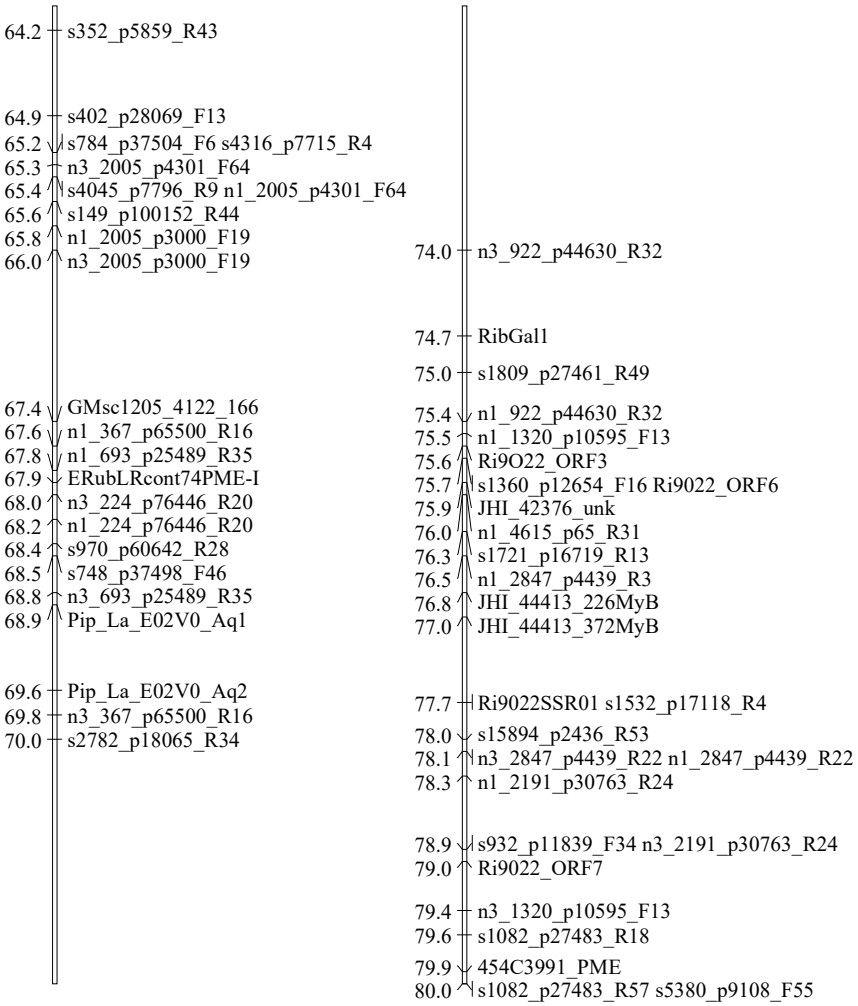
CONTINUE NEXT PAGE



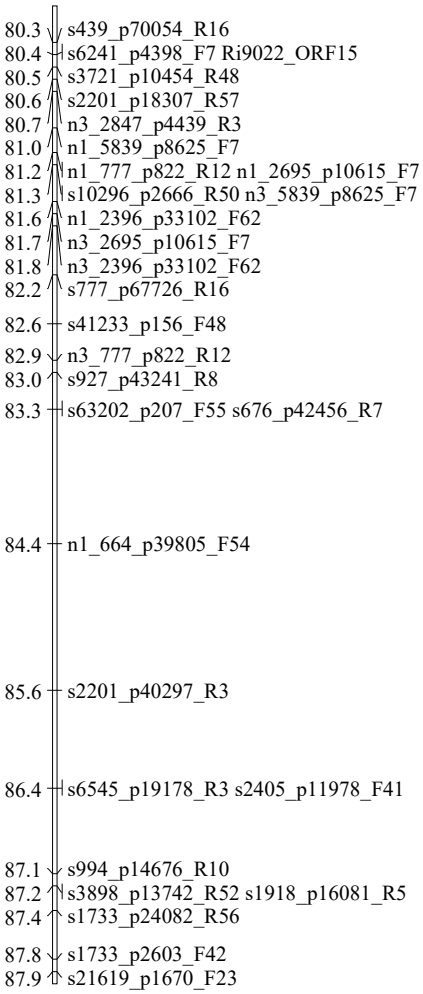
CONTINUE NEXT PAGE



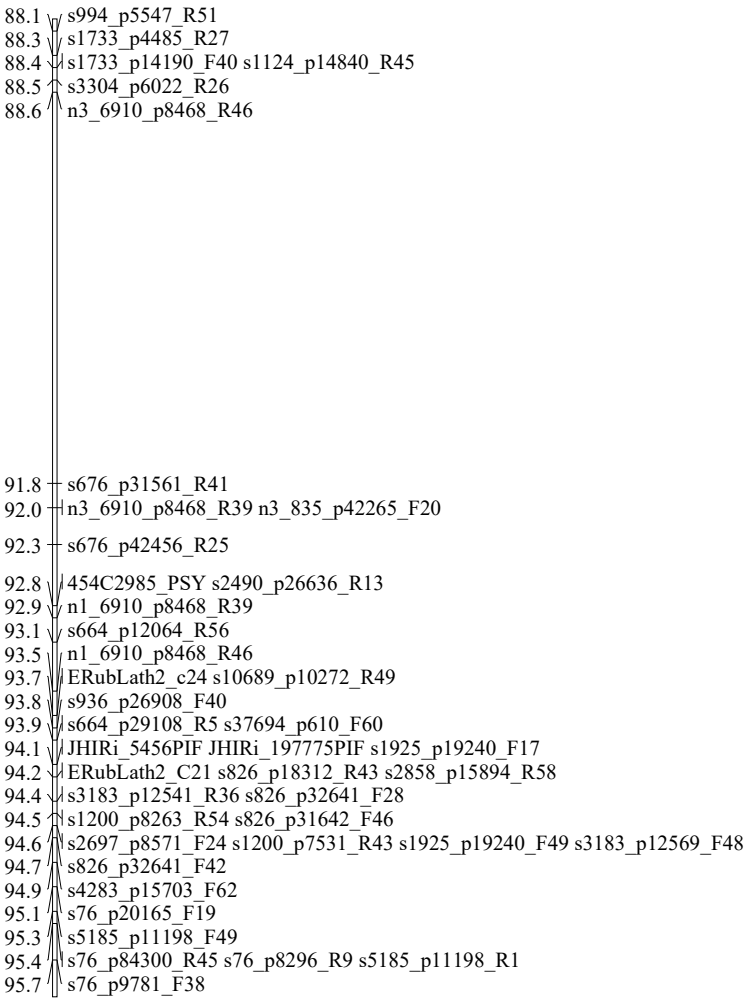
CONTINUE NEXT PAGE



CONTINUE NEXT PAGE

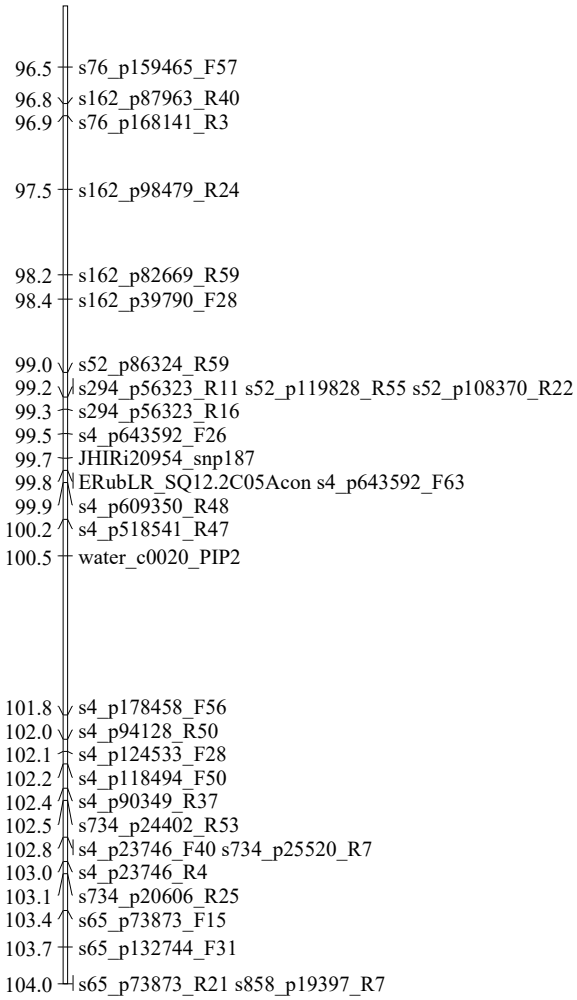


CONTINUE NEXT PAGE

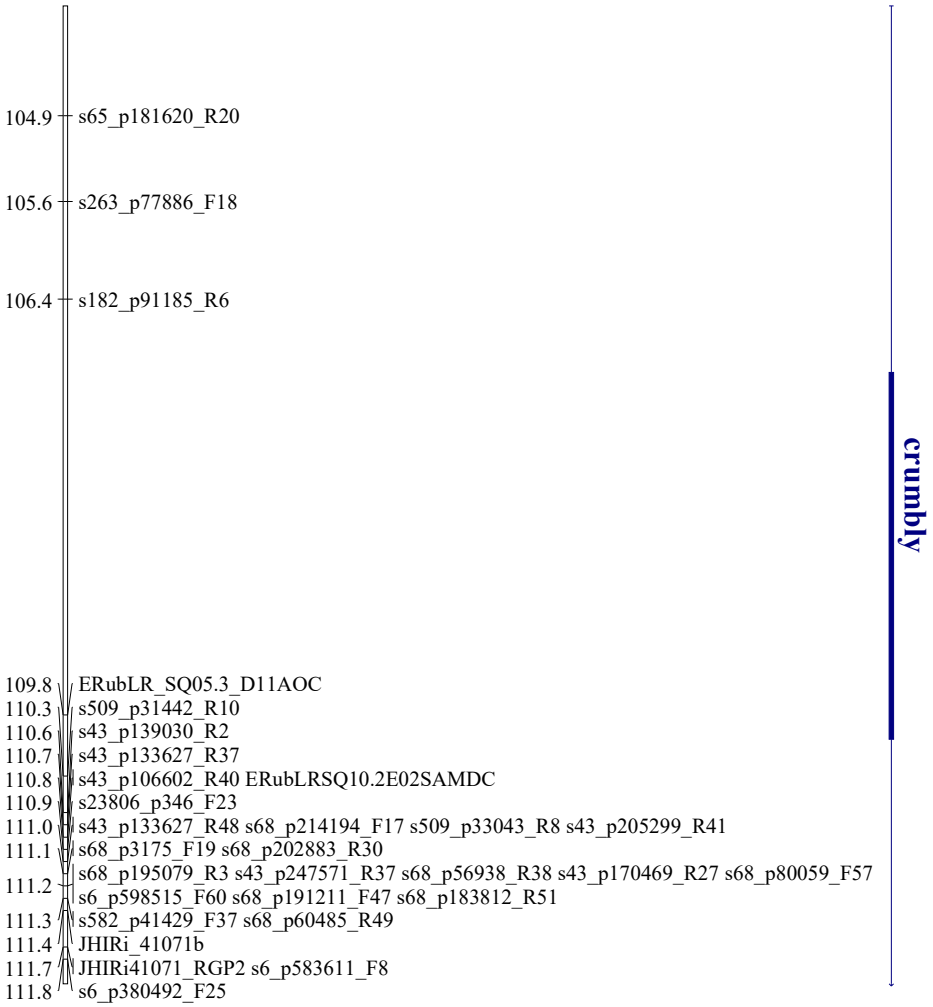


CONTINUE NEXT PAGE

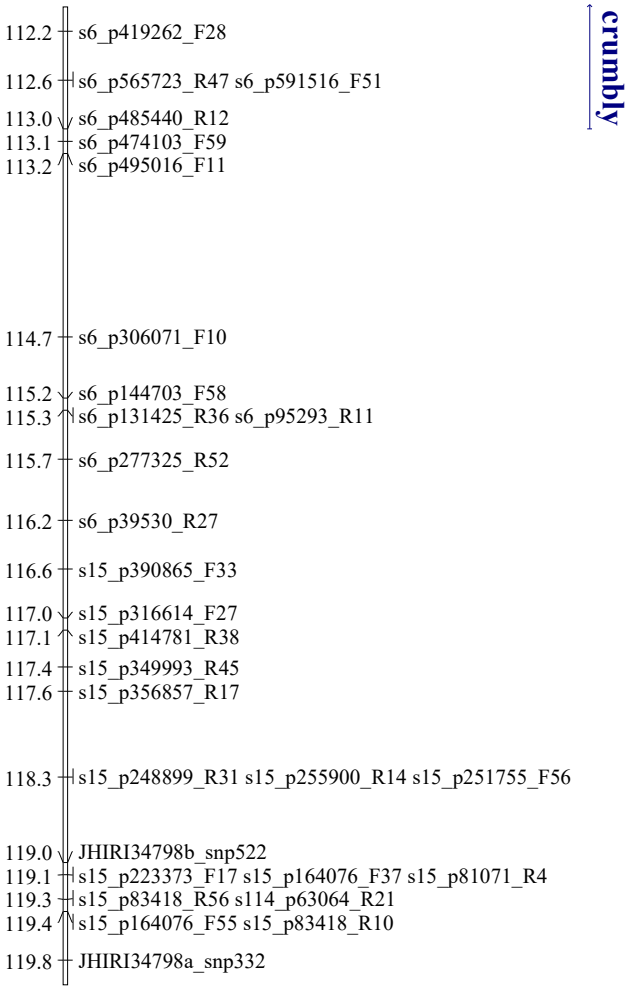
s [93]



CONTINUE NEXT PAGE



CONTINUE NEXT PAGE



CONTINUE NEXT PAGE

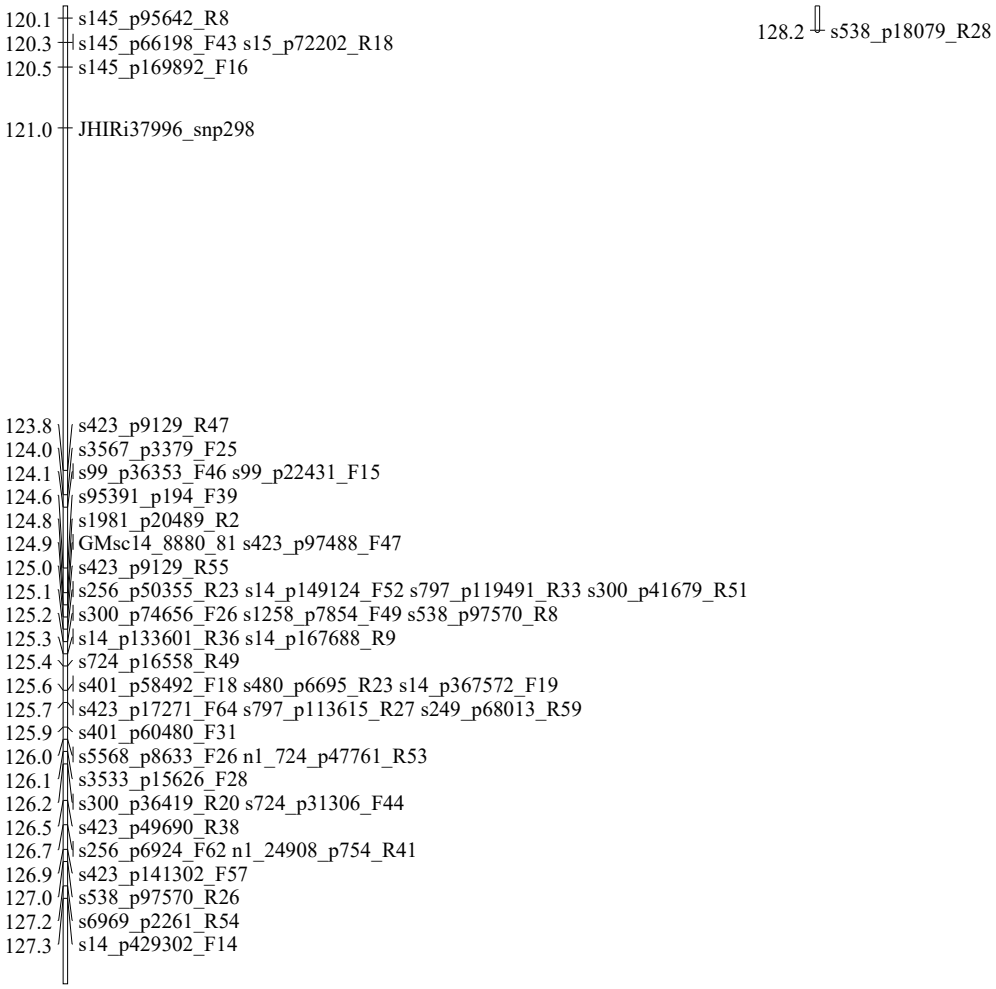


Figure A.4.2: **Lingage group 3 (LG3) full markers map and the two 'crumbly' QTLs.**